5-2012

Integrating Remote Sensing and Ecosystem Models for Terrestrial Vegetation Analysis: Phenology, Biomass, and Stand Age

Gong Zhang
Utah State University

Recommended Citation

Zhang, Gong, "Integrating Remote Sensing and Ecosystem Models for Terrestrial Vegetation Analysis: Phenology, Biomass, and Stand Age" (2012). *All Graduate Theses and Dissertations*. 1316.
https://digitalcommons.usu.edu/etd/1316

INTEGRATING REMOTE SENSING AND ECOSYSTEM MODELS FOR TERRESTRIAL VEGETATION ANALYSIS: PHENOLOGY, BIOMASS, AND STAND AGE

by

Gong Zhang

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Ecology

Approved:

Dr. Michael A. White Major Professor	Dr. Robert Gillies Committee Member
Dr. R. Douglas Ramsey Committee Member	Dr. Ronald J. Ryel Committee Member
Dr. Christopher Neale Committee Member	Dr. Mark R. McLellan Vice President for Research and Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2012

ABSTRACT

Integrating Remote Sensing and Ecosystem Models for Terrestrial Vegetation Analysis: Phenology, Biomass, and Stand Age

by

Gong Zhang, Doctor of Philosophy

Utah State University, 2012

Major Professor: Dr. Michael A. White
Department: Watershed Sciences

Terrestrial vegetation plays an important role in global carbon cycling and climate change by assimilating carbon into biomass during the growing season and releasing it due to natural or anthropogenic disturbances. Remote sensing and ecosystem models can help us extend our studies of vegetation phenology, aboveground biomass, and disturbances from field sites to regional or global scales. Nonetheless, remote sensing-derived variables may differ in fundamental and important ways from ground measurements. With the growth of remote sensing as a key tool in geoscience research, comparisons to ground data and intercomparisons among satellite products are needed. Here I conduct three separate but related analyses and show promising comparisons of key ecosystem states and processes derived from remote sensing and theoretical modeling to those observed on the ground. First, I show that the Moderate Resolution Imaging Spectroradiometer (MODIS) greenup product is significantly correlated with the earliest ground phenology event for North America. Spring greenup indices from different satellites demonstrate similar variability along latitudes, but the number of ground phenology observations in summer, fall, and winter is too limited to interpret the remote sensing-derived phenology products. Second, I estimate aboveground biomass (AGB) for California and show that it agrees with inventory-based regional biomass assessments. In this approach, I present a new remote sensing-based approach for mapping live forest AGB based on a simple parametric model that combines high-resolution estimates of Leaf Area Index derived from Landsat and canopy maximum height from the space-borne Geoscience Laser Altimeter System (GLAS) sensor. Third, I built a theoretical model to estimate stand age in primary forests by coupling a carbon accumulation function to the probability density of disturbance occurrences, and then ran the model with satellite-derived AGB and net primary production. The validated remote sensing data, integrated with ecosystem models, are particularly useful for large-region vegetation research in areas with sparse field measurements, and will help us to explore the long-term vegetation dynamics.

(123 pages)

PUBLIC ABSTRACT

Integrating Remote Sensing and Ecosystem Models for Terrestrial Vegetation Analysis: Phenology, Biomass, and Stand Age

by

Gong Zhang, Doctor of Philosophy

Utah State University, 2012

Major Professor: Dr. Michael A. White
Department: Watershed Sciences

Terrestrial vegetation, especially forest, stores plenty of carbon from carbon dioxide in atmosphere into biomass by photosynthesis, and serves as an important carbon sink in global carbon cycle. Meanwhile the carbon stored in live biomass of terrestrial vegetation is released due to natural or anthropogenic disturbances, such as wild fire, hurricane, and urbanization. The phenology, representing the seasonal change of vegetation, is a sensitive indicator for carbon fluxes via vegetation. Remotely sensed imagery from sensors abroad on satellite can help us extend our study of terrestrial vegetation from field sites to regional or global scales. Researchers have employed remote sensing data to monitor phenology change, forest biomass, and fire activities. Nonetheless, remote sensing-derived variables may be different in definitions and important ways from ground measurements. The more remote sensing data is used in geoscience research, the more validation with ground data and intercomparison among satellites are needed. Here I conduct three separate but related analyses and show comparison of key ecosystem states and processes derived from remote sensing and ecosystem models to those observed on the ground. First, I find that the greenup from Moderate Resolution Imaging Spectroradiometer (MODIS) product agrees with the earliest phenology event measured in various sites crossing

North America. The greenup indices from different satellites also show the similar latitude trend. However, the number of ground phenology observations in summer, fall, and winter is limited to validate the remote sensing derived phenology products. Second, I generate California forest aboveground biomass maps from Landsat imagery and space-borne Lidar data. I test my model over the forested areas of California with various climatic and land-use conditions and I show that my AGB estimates are comparable to inventory estimates. After the validation and intercomparison of those variables, an exciting opportunity lies in integrating these remote sensing derived data together by ecosystem model. I build a ecosystem model by linking the remote sensing-derived carbon accumulation rate, carbon storage and disturbance together, and produce the first global forest stand age map. My model with integrating remote sensing-derived data is particularly useful for vegetation areas with underrepresented field measurements, and will help us to explore the long-term vegetation history.

For my family…

ACKNOWLEDGMENTS

There have been many who have offered me great support as I have endeavored to do this dissertation. I would not have even close to obtain a Ph.D. without the support and encouragements from them.

I would like to thank my advisor Dr. Michael A. White, for getting me started on this path, and keeping his enthusiastic mentors through the course of my studies. I would like to thank my committee members for their patience and their interest in my research.

I would like to acknowledge on-going support from the NASA Ames Ecological Forecasting Lab and throughout the research phase of this dissertation. Dr. Ramakrishna Nemani, as leader of the Laboratory, provided direction, motivation, and support to reach the research goals I set. Special thanks go to Dr. Sangram Ganguly and Dr. Weile Wang, this dissertation obviously could not exist without their efforts. Colleagues in the Laboratory, Dr. Forrest Melton, Dr. Cristina Milesi, Dr. Petr Votava, Dr. Hirofumi Hashimoto, Dr. Jennifer Dungan, Mr. Andrew Michaelis and Mr. Sam Hiatt are all provided me many valuable advice and support.

I am extremely grateful to College of Natural Resources for awarding me the S.J. Quinney Ph.D. fellowship, and wish to extend special thanks to our department staff assistant for graduate students, Brian Bailey, for his great help to me through the maze of paperwork and deadlines.

I would like to thank my Loving, Deng Ding, supported me through hard times.

Finally, I must extend deserved recognition to my parents. Mum and Dad, you have support me through everything I've ever done. I am entirely dependent on your advice and patience.

Gong Zhang

CONTENTS

Page

CHAPTER

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

1.1 Background

The terrestrial ecosystem carbon exchange with the atmosphere is one of the most important components of the global carbon cycle, and can partially offset the course of anthropogenic greenhouse gas emissions (Canadell et al., 2007; IPCC, 2007). The atmosphere-ecosystem exchange of carbon also has a profound impact on climatic conditions at the Earth's surface (Sellers et al., 1996), ultimately driving various climatological processes at regional and global scales, like evapotranspiration, precipitation, and absorbed radiation (Dirmeyer, 1994). Therefore, accurate estimation of carbon fluxes and pools in terrestrial ecosystems in the context of global change is a challenging and important task.

Previous studies (e.g., Baldocchi and Vogel, 1996; Falge et al., 2002; Hanson et al., 2004; Reich et al. 1995) have shown that the net ecosystem exchange of carbon (NEE) from subtropical to boreal ecosystems varies seasonally, and the seasonal variability of NEE is closely related to vegetation phenology, the seasonal dynamic of vegetation. Phenophases, which refer to stages in the timing of plant development and physiological status, influence seasonal patterns of net carbon flux (White et al., 1999; Cramer et al., 1999). Vegetation in temperate regions is typically a net source of carbon during the period of plant dormancy and a net sink during the period of active growth. The seasonal change in NEE affects the concentration of atmospheric CO2 at local to hemispheric scales (e.g., Keeling et al. 1996; Murayama et al. 2003).

Various methods have been employed to explore vegetation phenology. Ground plant phenology events are measured at many ground observation sites to quantify plant seasonal change and to provide data for phenology and ecosystem modeling studies. But ground measurements are usually limited in terms of sites the number of observed species, thus restricting the application of ground phenology data for regional studies. With the virtue of high spatio-temporal resolution, remote sensing presents a promising alternative for large area phenology research (Schwartz, 2003). The beginning and end of the growing season, as well as other development periods of growth and reproduction (e.g. Zhang et al., 2003), may be extracted from time series of satellite vegetation indices. Yet identifying phenological stages for individual plants or for particular stages such as blooming remains problematic for remote sensing. Linking long-term ground measurements of plant phenology with remote sensing-derived land surface phenology products remains a fertile research field critical for scaling phenology from individual plants to canopies, regions and the globe.

Carbon storage in terrestrial ecosystems dramatically exceeds annual carbon fluxes, and changes in carbon storage has the potential to rapidly affect atmospheric CO2 content. Monitoring and preserving the carbon stored in terrestrial ecosystems can therefore help to limit ongoing climate change induced by greenhouse gasses (GHGs) or, as in the case of reducing emissions from deforestation and degradation (REDD), may actually lead to carbon storage (Miles and Kapos, 2008). Mapping and monitoring carbon stocks that change dynamically over time require spatially explicit estimates of forest cover (Saatchi et al., 2011; Baccini et al., 2008), and estimates of long-term carbon uptake in forests. A direct measurable property of carbon stored in forests is the forest biomass, in particular, the aboveground portion of the total biomass carbon (Gurney and Raymond, 2008).

The U.S. Forest Service Inventory and Analysis Program (FIA) provides the most extensive and accurate field observations of forest biomass and disturbance history for the U.S. and thus has been a key information source in many carbon studies (Keith et al., 2010). In many nations and regions, however, a detailed ground inventory is not available since it requires extensive labor, funding and materials for ground measurements. To date, several studies have provided an extensive knowledge base for estimating forest aboveground biomass (AGB). Currently, research on AGB is progressing rapidly with advancements in space-based modeling of the 3-D canopy structure, combined with canopy reflectance measured by passive optical sensors and radar backscatter (Lefsky et al., 1999; Drake et al., 2002; Asner et al., 2010; Saatchi et al., 2011). But most studies are either detailed assessments at local scales or large region estimates from coarse resolution remote sensing data. Few studies have looked at estimating high-resolution biomass density from finer resolution sensors such as Landsat or other commercial satellites over continental-to-regional and local scales.

Disturbances on ecosystems are an important regulator of the states and fluxes of ecosystem carbon, and can quickly transform forests from carbon sinks to carbon sources (Turner et al. 1993). In addition, carbon storage and fluxes in terrestrial ecosystems are linked with forest stand age under given disturbance regimes. Severe disturbances acting over large regions are recorded in forest inventories, but individual disturbance events are easily misinterpreted due to their temporal and spatial irregularity. Here, as for phenological research, remote sensing can provide an alternative method to pick up such disturbances from local to global scales during the past several decades. Yet disturbance and recovery cycles in forests may last decades to hundreds years, mandating theoretical approaches to combine the recent observational record with longer-term processes.

1.2 Objectives

My goal is to link remote sensing data with ecosystem models to provide a broad analysis of vegetation dynamic and conditions. In particular, I have the following objectives in the areas of phenology, AGB and stand age:

1) Compare the Moderate Resolution Imaging Spectroradiometer (MODIS) phenology products of multi-phenophase information with plant phenology events during spring,

summer and fall. Combine the ground plant phenology records from various observational networks and evaluate the relationship between the most appropriate plant phenology events and remotely sensed land surface phenology. Compare the MODIS phenology products with other state-of-the-art phenology products derived from different algorithm and sensors.

2) Develop a method to convert tree height to AGB model based on spaceborne Lidar metrics and a Landsat vegetation index, and apply the approach for a large region. Compare my AGB map and other existing satellite-derived AGB density estimates with available FIA inventory data at an aggregated regional level. Address spatial scaling issues arising from comparing satellite estimates of AGB density across various resolutions and with forest inventory plot level measurements.

3) Build a theoretical framework of disturbance regime based on AGB and net primary production (NPP). Then try to estimate long-term disturbance regimes in primary forests using current remote sensing data, and extract forest turnover rate and stand age.

1.3 Motivation

As such, all objectives rely on remote sensing data with global coverage at low cost. Thematically, all of my research is related to the vegetation carbon cycle and may ultimately prove useful in international efforts to consistently and transparently measure, monitor, and verify terrestrial carbon sequestration.

References

Asner, G.P., Powell, G.V.N., Mascaro, J., Knapp, D.E., Clark, J.K., Jacobson, J., Kennedy-Bowdoin, T., Balaji, A., Paez-Acosta, G., Victoria, E., Secada, L., Valqui, M., Hughes, R.F., 2010. High-resolution forest carbon stocks and emissions in the Amazon. *Proceedings of the National Academy of Sciences* 107, 16738 –16742.

Baccini, A., Laporte, N., Goetz, S.J., Sun, M., Dong, H., 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. *Environmental Research Letters* 3, 045011.

Baldocchi, D.D., Vogel, C.A., 1996. Energy and CO2 flux densities above and below a temperate broad-leaved forest and a boreal pine forest. *Tree Physiology* 16, 5–16.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., 2001. FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society* 82, 2415–2434.

Canadell, J.G., Le Quéré, C., Raupach, M.R., Field, C.B., Buitenhuis, E.T., Ciais, P., Conway, T.J., Gillett, N.P., Houghton, R.A., Marland, G., 2007. Contributions to accelerating atmospheric CO2 growth from economic activity, carbon intensity, and efficiency of natural sinks. *Proceedings of the National Academy of Sciences* 104, 18866.

Cramer, W., Kicklighter, D.W., Bondeau, A., Iii, B.M., Churkina, G., Nemry, B., Ruimy, A., Schloss, A.L., 1999. Comparing global models of terrestrial net primary productivity (NPP): overview and key results. *Global change biology* 5, 1–15.

Dirmeyer, P.A., 1994. Vegetation stress as a feedback mechanism in midlatitude drought. *Journal of climate* 7, 1463–1483.

Drake, J.B., Dubayah, R.O., Knox, R.G., Clark, D.B., Blair, J.., 2002. Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest. **Remote Sensing of Environment** 81, 378–392.

Falge, E., Baldocchi, D., Tenhunen, J., Aubinet, M., Bakwin, P., Berbigier, P., Bernhofer, C., Burba, G., Clement, R., Davis, K.J., 2002. Seasonality of ecosystem respiration and gross primary production as derived from FLUXNET measurements. *Agricultural and Forest Meteorology* 113, 53–74.

Gurney, K.R., Raymond, L., 2008. Targeting deforestation rates in climate change policy: a "Preservation Pathway" approach. *Carbon Balance Manag* 3, 2.

Hanson, P.J., Amthor, J.S., Wullschleger, S.D., Wilson, K.B., Grant, R.F., Hartley, A., Hui, D., Hunt, J., Johnson, D.W., Kimball, J.S., 2004. Oak forest carbon and water simulations: model intercomparisons and evaluations against independent data. *Ecological Monographs* 74, 443–489.

IPCC, C.C., 2007. *The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA 996, 2007.

Keeling, C.D., Chin, J.F.S., Whorf, T.P., 1996. Increased activity of northern vegetation inferred from atmospheric CO2 measurements. *Nature* 382, 146–149.

Keith, H., Mackey, B., Berry, S., Lindenmayer, D., Gibbons, P., 2010. Estimating carbon carrying capacity in natural forest ecosystems across heterogeneous landscapes: addressing sources of error. *Global Change Biology* 16, 2971–2989.

Lefsky, M.A., Harding, D., Cohen, W.., Parker, G., Shugart, H.., 1999. Surface Lidar Remote Sensing of Basal Area and Biomass in Deciduous Forests of Eastern Maryland, USA. *Remote Sensing of Environment* 67, 83–98.

Miles, L., Kapos, V., 2008. Reducing Greenhouse Gas Emissions from Deforestation and Forest Degradation: Global Land-Use Implications. *Science* 320, 1454 –1455.

Murayama, S., Saigusa, N., Chan, D., Yamamoto, S., Kondo, H., Eguchi, Y., 2003. Temporal variations of atmospheric CO2 concentration in a temperate deciduous forest in central Japan. *Tellus B* 55, 232–243.

Reich, P.B., 1995. Phenology of tropical forests: patterns, causes, and consequences. *Canadian Journal of Botany* 73, 164–174.

Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M., Mitchard, E.T.A., Salas, W., Zutta, B.R., Buermann, W., Lewis, S.L., Hagen, S., Petrova, S., White, L., Silman, M., Morel, A., 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences* 108, 9899 –9904.

Schwartz, M.D., 2003. *Phenology: an integrative environmental science*. Springer-Verlag, Berlin.

Sellers, P.J., Randall, D.A., Collatz, G.J., Berry, J.A., Field, C.B., Dazlich, D.A., Zhang, C., Collelo, G.D., Bounoua, L., 1996. A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation. *Journal of climate* 9, 676–705.

Turner, M.G., Romme, W.H., Gardner, R.H., O'Neill, R.V., Kratz, T.K., 1993. A revised concept of landscape equilibrium: disturbance and stability on scaled landscapes. *Landscape Ecology* 8, 213–227.

White, M.A., Running, S.W., Thornton, P.E., 1999. The impact of growing-season length variability on carbon assimilation and evapotranspiration over 88 years in the eastern US deciduous forest. *International Journal of Biometeorology* 42, 139–145.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.., Gao, F., Reed, B.C., Huete, A., 2003. Monitoring vegetation phenology using MODIS. *Remote sensing of environment* 84, 471–475.

CHAPTER 2

EVALUATION OF THE MODIS PHENOLOGY PRODUCT WITH GROUND MEASUREMENTS AND INTERCOMPARISON WITH EXISTING PHENOLOGY PRODUCTS FROM THE MERIS AND AVHRR SENSORS FOR NORTH AMERICA

Abstract

Vegetation phenology can serve as a sensitive diagnostic of ecosystem response to global change. Long-term ground measurements of plant phenology have been instrumental in our understanding of local patterns of seasonality with changing climate, but ground measurements are focused on particular species or individual plants at specific sites across different observational networks. Their relationship to landscape change is therefore often unclear. Conversely, land surface phenology derived from satellite remote sensing can monitor seasonal change in ecosystems across large areas, but suffers from unknown relationships to plant phenology. The moderate-resolution imaging spectroradiometer (MODIS) Global Land Cover Dynamics Product (MOD12Q2) provides estimates of seasonal and interannual dynamics in vegetation phenology (metrics related to start of growing season, maturity, senescence and dormancy) at a 500 m resolution. The MOD12Q2 has been validated previously with ground plant phenology measurements at a few representative sites, but an comparison of MODIS and ground phenology at a continental scale is absent. Here, I compared the MODIS phenology product with ground phenology data from numerous sites and networks across North America. On average, the MODIS onset of greenup was 3.1 day earlier than the earliest plant spring phenological events, implying that MODIS can reliably capture vegetation greenup across North America. The results also suggest that heterogeneous ground phenology data should not be universally compared to remote sensing data. Instead, comparisons of ground phenology data at intensively measured sites should be used to first identify the plant stage that most closely

corresponds to remotely-sensed estimates. Further, intercomparison of MODIS phenological phases in summer, fall, and winter with the corresponding plant phenological events suggests that the MODIS algorithm is robust in capturing different phases of vegetation growing cycle, although with the exception of spring, ground measurements are limited. I then compared MOD12Q2 to two additional greenup products: (1) greenup date calculated based on the midpoint method with the Advanced Very High Resolution Radiometer (AVHRR) data; and (2) onset of greenup detected as the time series inflection point from the Medium Resolution Imaging Spectrometer (MERIS) Terrestrial Chlorophyll Index (MTCI). Results show a broad agreement in zonal averages of green-up dates with the MODIS and MTCI products, while AVHRR was 10 to 20 days late in middle and low latitude regions. Overall, the moderate resolution remotely sensed phenology products capture the spatial pattern of vegetation phenology across different landscapes in North America.

2.1 Introduction

Vegetation phenology represents the seasonal dynamics of vegetation and their connection to climate, and varies with species and their geographical distribution (Schwartz, 2003). Temperature plays a significant role in spring phenology such that a warming spring is attributed to an earlier green-up of temperate vegetation (Badeck et al., 2004; Chmielewski & Rötzer, 2001; Schwartz et al., 2006), although some studies suggest that colder temperatures in winter are required to meet the chilling requirements for greenup of forests and crops (Cesaraccio et al., 2004; Zhang et al., 2007). Fall phenology (e.g. leaf senescence) is related to photoperiod where shorter day lengths induce dormancy (White et al., 1997). Other abiotic factors, such as precipitation, soil, and topographic variables, also influence plant phenology in different ways. For instance, in the tropics and subtropics, the start of the rainy season controls the photochemical capacity in trees (Ishida et al., 2006). Therefore, vegetation phenology is now widely used as an ecological indicator to understand and quantify the impact of global changes on ecosystems.

Vegetation phenology also has a strong impact on seasonal and interannual variability in ecosystem carbon, water and energy cycling (Keeling et al., 1996; Körner & Basler, 2010; Myneni et al., 1997; Peñuelas & Filella, 2009). Growing-season length, the period from the start of growing season (SOS) to the end of growing season (EOS), is related to the timing of vegetation photosynthesis and heterotrophic respiration (Euskirchen et al., 2006; Piao et al., 2008). In addition, the overall vegetation phenological cycle can be divided into several distinct phenophases, each of which is related to plant development, morphological and physiological processes (Morisette et al., 2008). Characteristics of plant photosynthesis and, in turn, the amounts of carbon exchange vary with phenological stages, like higher photosynthetic capacity during leaf expansion as compared to leaf senescence (Wilson et al., 2000). Accurate knowledge of the timing in plant phenophases is therefore required for observing and understanding intra- and inter-annual variability in terrestrial ecosystem processes.

As any local community is composed of a number of plant species, vegetation phenology for a plant community is distinct from the phenology of individual species inside a community (Forrest et al., 2010). In essence, vegetation phenology is the combination of plant phenology dynamics for all species coexisting in a community. Plant phenology, the timing of phenological events on individual plants, provides a biological indicator of historical and current climate change and its impact on vegetation (Chuine et al., 2004; Schwartz & Reiter, 2000). Species-level phenological observations have been recorded by individuals and organizations in North America for several decades, although recording frequency and extent is irregular (Schwartz, 2003). Ground phenological observation usually focus on phenological events of individuals, involving the leaf, flower and seed that can be easily observed and measured by amateur observers without special skills or tools.

The leaf is the core organ where photosynthesis takes place, so “leaf out” -- or initial leaf expansion -- is regarded as a signal of SOS. Leaf senescence is another key stage in plant growth as it indicates the decline of photosynthesis and the end of the growing season. Besides the events related to leaves, vegetation phenological events occur in other parts of a plant as documented by in situ observations, such as blooming and seeding that represent a reproductive cycle as compared to a growth cycle. Blooming, which occurs even before leaf-out for some species, is considered as an early signal of plant development. Due to various competition and reproduction strategies, however, blooming for specific species may be far posterior to leaf-out and the start of photosynthesis, and may even occur in different phases of the calendar year. The variety of events occurring during the growing season may provide useful cues in identifying specific events, but they also muddle attempts to identify a dominant cardinal date for SOS or EOS.

A few species with visually observable events act as standard biological indicators for environmental change studies in large-scale operational networks, for e.g., the North American Lilac-Honeysuckle Observation Network (Schwartz, 1994). Because people tend to select common species whose phenological events are readily and inexpensively identified, phenological greenup data are biased towards events like lilac blossom as opposed to leaf flushing of evergreen plants. Phenological networks thus represent a confined set of species with a wide presence across different bio-climatic zones and provide a spatially extensive platform for robust and consistent measurements. But the observations based on these few common species may not completely illustrate the phenological patterns in large regions across various local vegetation communities.

Phenological observational sites, operated by professional staff, for example the Long Term Ecological Research Network (LTER, http://globalmonitoring.sdstate.edu/LTER-phenology/), measure a number of phenological events amongst various representative species in the local community. Such extensive observations of species at one site contain abundant phenological information for a whole community, which conversely limits the number of locations possible due to high labor and knowledge costs. The ground phenology records from species extensive and spatial extensive networks each have their advantages, but it is not clear to what extent of these two types of ground phenology datasets characterize vegetation phenology and landscape seasonal pattern.

The term land surface phenology (LSP), representing an integrated surface phenology estimate from remotely sensed observations, is used to explore the physical connection between ground observed phenological events and satellite derived estimates (De Beurs & Henebry, 2004; Ganguly et al., 2010). The signal measured by a satellite sensor, in most cases, is a convolved reflectance intensity value that results from the interaction between atmosphere, vegetation, snow, soil, and other objects above or on the land surface. Thus, LSP usually represents mixed information from various species in the local vegetation community as well as other objects in a pixel, and for this reason the LSP is also used as a reference to detect the timing of snowmelt, soil thaw and freeze (Kimball et al., 2004).

Of course, it is still difficult if not impossible to identify species or individual plants at the native spatial resolution of moderate-to coarse resolution satellite sensors, and LSP estimates must be taken a broad indication of community phenological development. Several studies have proposed methodologies to detect events such as SOS or EOS from time-series of remote sensing data, predominantly using vegetation indices (VIs) as a proxy for modeling the seasonal behavior of vegetation activity (e.g. Dash et al., 2010; Liang & Schwartz, 2009; Reed et al., 1994; Thompson & Wehmanen, 1979; Zhang et al., 2003;). VIs, such as the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI), are formulated to measure photosynthetically active vegetation (Huete et al., 2002), however, the VI time-series are noisy due to effects introduced by mixed-pixel behavior and atmospheric contamination. Multiple distinct phenologies within a single pixel also make it difficult to determine events like SOS or EOS from a gradual change in a VI. Further, selection of a technique for LSP detection may influence the LSP results. There has been a systematic evolution of methodologies in detecting the timing of greenup from VI time-series (White et al., 2009) but the majority of the algorithms can be broadly classified into four groups – (1) global thresholds - the date when the magnitude of the VI time series exceeds a “unique threshold” ; (2) local thresholds – instead of having a “unique threshold” like the global threshold method, this method uses a tuned value of the threshold, for e.g., the midpoint method where the threshold is between the maximum and minimum of the VI time series; (3) conceptual-mathematical, modeling the phenological evolution through mathematical fitting functions, for e.g., the logistic growth function as in MOD12Q2 product (Ganguly et al., 2010); and (4) hybrid, combining two or more of the above methods (e.g. the Gaussian method in Myneni et al., 1997).

Spectral reflectances from different sensors show characteristic bias due to differences in: (a) purity of the pixel containing a target (mixture vs. pure classes); (b) spectral differences in the wavelength bandwidth; (c) viewing and illumination geometry; (d) pre-calibration and/or atmospheric correction procedures if any; and (e) geolocation uncertainties. Therefore, the time series of VIs calculated from the different algorithms will differ in magnitude and phase for the same spatial grid; e.g. , the MODIS VI time series at a moderate spatial resolution displayed greater sensitivity of seasonal change compared to the Advanced Very High Resolution Radiometer (AVHRR) VI at a coarse spatial resolution (Huete et al., 2002). Thus LSP metrics derived from time series of VIs from different sensors may have inconsistencies and hence unknown uncertainties in phenological estimates.

Both LSP and ground plant phenology are related to seasonal changes in terrestrial vegetation, although they differ in terms of definition, materials, methodology, and scale – ideally they provide complementary information. The straightforward ground plant phenological measurements on individual plants lead to a dilemma in the balance between species and spatial representativeness, while the large spatial coverage of remote sensing data enables LSP estimation at regional-to–global scales and may not capture the species level information. It is challenging, but valuable to attempt to link LSP and ground plant phenology given their different sets of strengths and weaknesses and different spatial scales of representation (Morisette et al. 2008). Such a linkage becomes crucial in order to measure the vegetation seasonal changes with the aid of satellite-derived metrics. Several studies have attempted to compare remote sensing-derived SOS with related ground phenological measurements (Liang et al., 2011; Schwartz & Hanes, 2010; White et al. 2009); however, there a re difficulties in directly relating AVHRR based LSP to species level phenological events as observed by ground measurements, such as point vs. pixel problem and different measurement protocols (White et al., 2009).

In this study, my goal is to understand the difference and agreement between ground phenological observations and LSP, in particular MCD12Q2, across North America. Specifically, I focus on two aspects:

(a) Direct comparison: Previous studies have compared satellite derived greenup metrics with ground measurements from representative sites (Ganguly et al., 2010; Hufkens et al., 2011; Zhang et al., 2006); however, comparisons of ground measured phenological events against LSP metrics like maturity, senescence and dormancy (Zhang et al., 2006) are rare. Here, I performed a comprehensive evaluation of the latest version of the MODIS derived phenology metrics using various ground observations across North America. I included ground measurements recording the onset of greenup, maturity, senescence and dormancy during the spring, summer and fall period. My specific aim was to elucidate the physical connection between species- and plant-specific phenological events and the MODIS LSP metrics, which has not been attempted at the continental scale.

(b) Indirect comparison: A previous study, using AVHRR measurements, has demonstrated the incongruities in SOS patterns due to differences in estimation algorithms (White et al., 2009). At regional- to continental-scales, however, intercomparison studies of greenup metrics derived from different satellite sensors are again rare (Fishard and Mustard, 2007; Liang et al., 2011). To this end, I made an indirect comparison of the greenup onset timings derived from three sensors - MODIS, AVHRR, and the Medium Resolution Imaging Spectrometer (MERIS). All the three sensors are distinct in terms of spatial resolution and spectral characteristics, so LSP products are based on different detection algorithms. Note that my aim here is to compare leading LSP estimation products, rather than the difference of detection algorithms (the subject of White et al., 2009).

The following sections describe the data used for this study, methodologies in extracting phenology metrics from ground observations and satellite measurements, and present the results from this intercomparison exercise.

2.2 Data and Methology

2.2.1 Ground observation phenology data

Unlike the previous study with only greenup records (White et al., 2009), I used a total of 7011 phenological events acquired from 19 networks and sites across Northern America from 2001 to 2006. I grouped events into three phenophases: greenup, maturity, and senescence (Table 2.1, Figure 2.1). These records are either spatially extensive or species extensive, and in addition include phenological events during summer and fall to evaluate the multi-phase satellite-derived phenological metrics in this study. Flowering, representing the reproductive timing, may occur before or after greenup, and over more extended period than leaf out (Lechowicz, 1995). The plants flowering in summer generally separate the reproductive growth phase from vegetative growth phase, so flowering time of these plans is irrelevant to greenup or SOS (Kudo et al., 2008). Ideally, it would be possible to screen the data by species and to include only those flowering events that are known to be contemporaneous with leafing. Unfortunately such species-specific data is not generally available, and I classified all blooming events into greenup due to the limited knowledge about species-specific phenology.

Table 2.1: Events recorded in phenology datasets. There are 13 phenological events observed; the number of *observations for each event* varies from 2 to 2,829 during 2001-2006, *individual* sites con*tain between* 1 to 413 locations and 2 to 535 species.

Development	**Events**	**Abbreviation**	**Total Records**	**Number of Locations**	**Number of Species**
Greenup	Leaf bud burst	LBB	459	90	80
	First leaf	FL	1287	162	273
	75% leaf expanding	75LE	49	1	17
	Full expansion	FE	10	1	2
	First flower bud	FLB	100	3	37
	First flower	FF	2829	351	483
	Middle blooming	MB	1513	238	85
Maturity	Seed maturity	SM	2	2	2
	Seed shedding	SS	150	1	17
Senescence	First leaf color change	FCC	445	2	20
	Leaf fall begin	LFB	13	2	2
	Last leaf color change	LCC	143	2	17
	Leaf fall complete	LFC	11	2	2

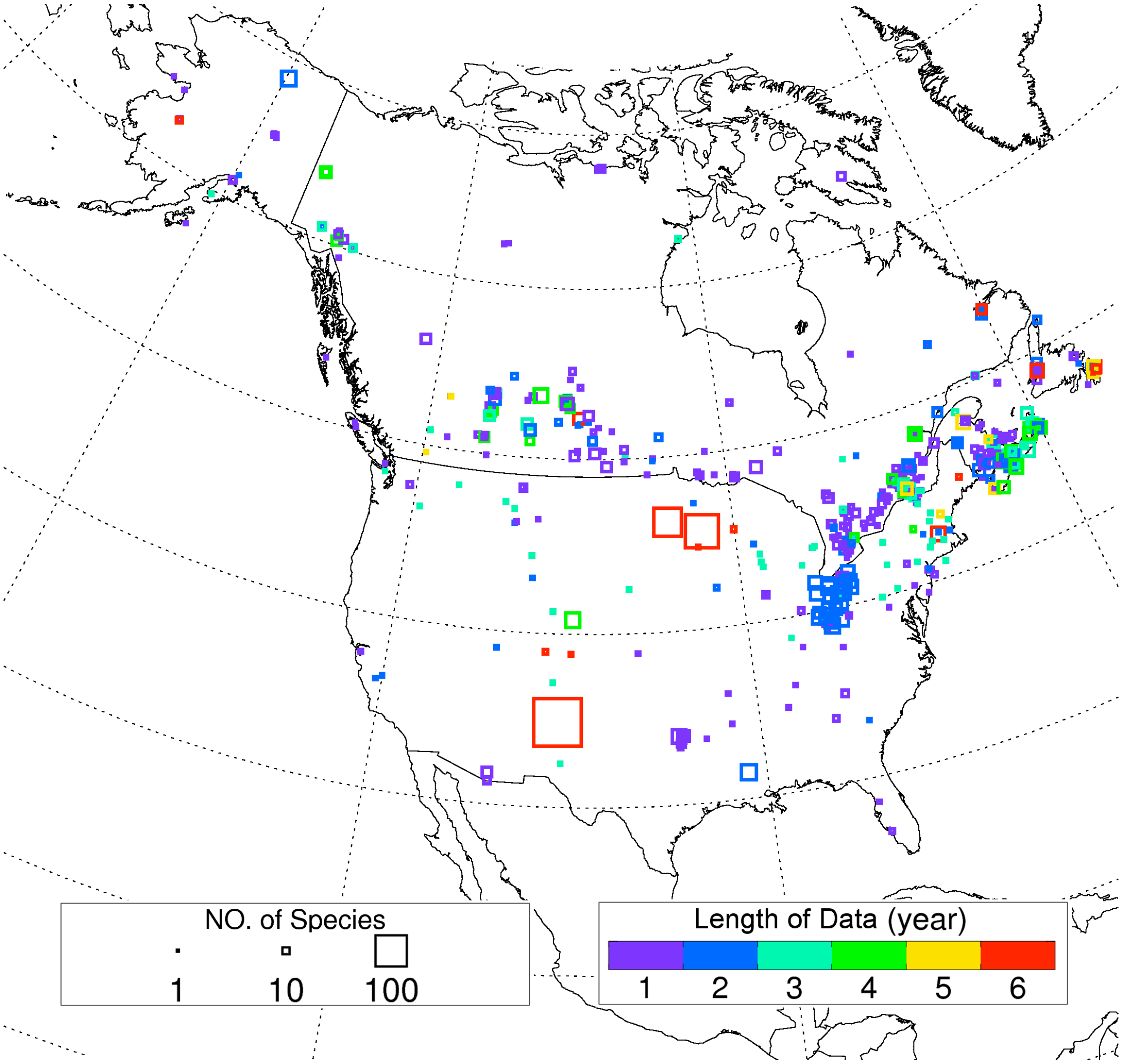


Figure 2.1: The location of ground measurements used. The duration of observation at each location is shown in color, and the number of observed species for each location is illustrated as box size.

To illustrate the phenology of a local community, I gathered the phenological events for a specific location that occurred at same year into one year-location cluster. I selected the earliest event from all greenup events within each year-location cluster to represent the onset of greenup rather than pooling the entire greenup phenophase record. Similarly, I selected the latest senescence events from the year-location clusters for the end of the growing season or onset of dormancy. Although these earliest/latest records may not capture the beginning or end of the growth curve, this approach provides a uniform approach for estimating phenological timing for species extensive sites in order to match the definition of LSP indices.

2.2.2 Land surface phenology data

Specific approaches have been developed to capture the timing of vegetation growth stages by utilizing satellite remote sensing derived products. MODIS sensors aboard the Terra and Aqua satellites offer a valuable data source to monitor regional-to-global scale land surface dynamics in the last decade at a spatial resolution of 500 m to 1 km. The MODIS vegetation index product suite provides both the NDVI and the Enhanced Vegetation Index (EVI) - the later incorporates the blue spectral band and is calculated as,

$$EVI = 2.5\frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + 6 \times \rho_{red} - 7.5 \times \rho_{blue} + 1} \qquad \textit{Eq. 2.1}$$

where (0.62-0.67μm), (0.84-0.88μm), and (0.46-0.48μm) are reflectances for narrow wavelength intervals as compared to AVHRR. The MODIS EVI can capture vegetation variations in high biomass areas and/or drought conditions, whereas the NDVI relies on the red band and tends to saturate in closed canopies (Huete et al., 2002). Based on EVI, the MODIS Land Cover Dynamics (MCD12Q2) product is produced to estimate the timing of vegetation phenology at global scales, and I used the 500 m refined Collection 5 (C5) MCD12Q2 product for this study (Ganguly et al. 2010). The MCD12Q2 product defines four phenophases in a typical plant development cycle: greenup, maturity, senescence, and dormancy. The product fits a logistic function to the EVI curve in one cycle, and the four phases are demarcated by calculating the maximum rate of change in curvature of the fitted logistic curve. The MODIS product estimates more plant phases than other products, which may be beneficial for estimating seasonal changes in photosynthesis and other plant activities (Sakamoto et al., 2005; Zhang, 2006).

To monitor long-term changes in phenological shifts, many studies have utilized the AVHRR products, especially the Normalized Difference Vegetation Index (NDVI). NDVI is computed from spectral reflectances as,

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \qquad \textit{Eq. 2.2}$$

where (~0.58-0.68μm) is the reflectance in red channel , and (~0.72-1.1μm) is the reflectance in NIR channel. The Global Inventory Modeling and Mapping Studies (GIMMS) dataset provides a l981 to present 16-day composited NDVI at global 8-km grids (Tucker et al., 2005), and is widely used for global and regional phenology estimates. In a prior intercomparison study between different methods based on NDVI from AVHRR, the pixel midpoint method based on the maximum and minimum of annual NDVI curve was found to be more closely related to ground plant phenology than other algorithmic methods (White et al., 2009). In t his study, I used the GIMMS data and the pixel midpoint method to calculate the start of growing season (Figure 2.2).

MERIS, onboard Envisat, is another valuable sensor for monitoring LSP at regional to global scales, since it has 15 programmable (2.5 nm-20 nm wide) wavebands in the region of 390 nm to 1040 nm and a spatial resolution of 300 m with a three-day repeat cycle. Because of MERIS's fine spectral and moderate spatial resolution, the MERIS Terrestrial Chlorophyll Index (MTCI) combines three MERIS visible and infrared bands (Dash & Curran, 2004) to provide information related to vegetation chlorophyll content and is calculated as,

$$MTCI = \frac{\rho_{Band10} - \rho_{Band9}}{\rho_{Band9} - \rho_{Band8}} \qquad \textit{Eq. 2.3}$$

where ρ_{Band8} , ρ_{Band9} , and ρ_{Band10} are the reflectances of bands 8 (681.25±7.5nm), 9 (708.75±10nm) and 10 (853.75±7.5nm), and respectively in the MERIS standard band setting. Dash et al (2010) developed a discrete Fourier transform approach to smooth MTCI 8-day time series dataset, and then identified the inflection points with maximum curvature. A change in derivative from positive to negative indicates a valley point (with the algorithm moving backwards in time), which in turn is identified as onset of greenup (Figure 2.2). This hybrid method for MERIS MTCI was found to be a reliable algorithm for measuring greenup of both natural vegetation and cropland areas with complex landscapes and multiple growth cycles (Dash et al., 2010). I used the MERIS MTCI greenup metrics for the conterminous USA (CONUS), which is available from 2003 to 2006 at 1 km spatial resolution.

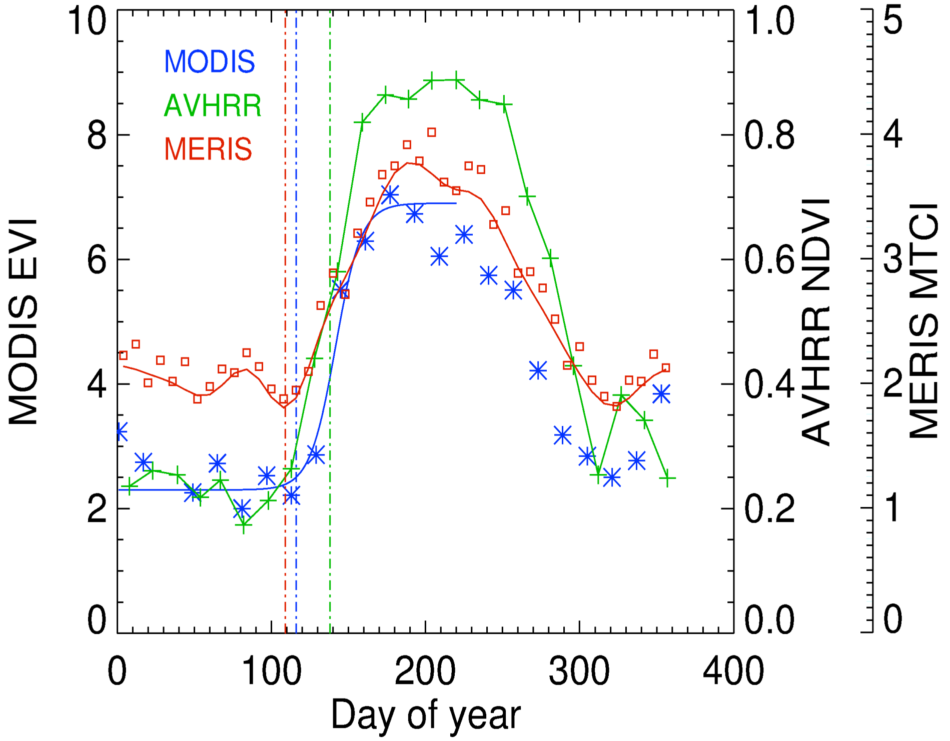


Figure 2.2: The LSP onset of greeup detection algorithms illustration for a representative forest pixel in California. Onset of greenup (vertical lines) for three methods and datasets: midpoint method with AVHRR NDVI linearly interpolated daily curve (green); MODIS EVI time series by logistic function fitting (blue), MERIS MTCI smoothed curve by discrete Fourier transform (red line).

2.2.3 Comparison

I compared the MODIS derived onset of greenup, maturity, senescence, and dormancy to ground phenological events from 2001 to 2006 (duration limited by the availability of MCD12Q2 at the time of research). I matched the timings of ground phenological events with MCD12Q2 phenophases extracted from the pixel of the native 500 m imagery nearest to locations of the ground records. If MCD12Q2 is missing in the corresponding pixel due to clouds or other effects, I discarded this pixel from the comparison (rather than using the value from a neighboring pixel). To quantify the comparison, I calculated the Pearson's correlation coefficients, bias, and root mean square error (RMSE) between MCD12Q2 and ground phenology. Instead of ordinary least squares (OLS) regression, I used orthogonal regression analysis (reduced major axis, RMA) to explore the relationship between LSP and observation (Berterretche et al., 2005), which is more appropriate in accounting simultaneously for the errors in both variables (Sokal and Rohlf, 1995).

I produced the AVHRR phenology metrics every year from 2001 to 2006, but the MTCI phenology production is only available from 2003 to 2006 for CONUS. To provide a more direct intercomparison among sensors, I averaged both the MODIS 500 m greenup data and MERIS MTCI 1 km greenup data to an 8 km spatial resolution to match the AVHRR products. I eliminated pixels with missing data in any products to assure unified standards in calculating the statistical metrics. I processed the intercomparison amongst MODIS, MTCI, and AVHRR phenology products for all available pixels in CONUS from 2003 to 2006. Further, due to the different definitions of greenup for the MODIS and AVHRR products, I also calculated their standard deviation for North America from 2001 to 2006 (to reflect interannual change).

2.3 Results and Discussion

Ground observations of different phenological events span early spring to late fall (Figure 2.3). Most ground phenological events, especially when classified into the three phenophases, occur around the expected MODIS events. Often the median date of each MODIS phenological event is within the interquartile range of the corresponding phenophases. Median “leaf bud burst” is close to onset of greenup, while median “75% leaf expending” and median “Full expansion” are near the end of greenup. Some ground events, however, like “First leaf” (1287 records) and “First flower” (2829 records) extend over a broader range (Figure 2.3). “First leaf” events may contain late leaf-out of drought-enduring plants in some arid/semi-arid sites, for example the Sevilleta LTER sites in New Mexico. “First flower” also has a large range, and its median value is later than day of year 200.

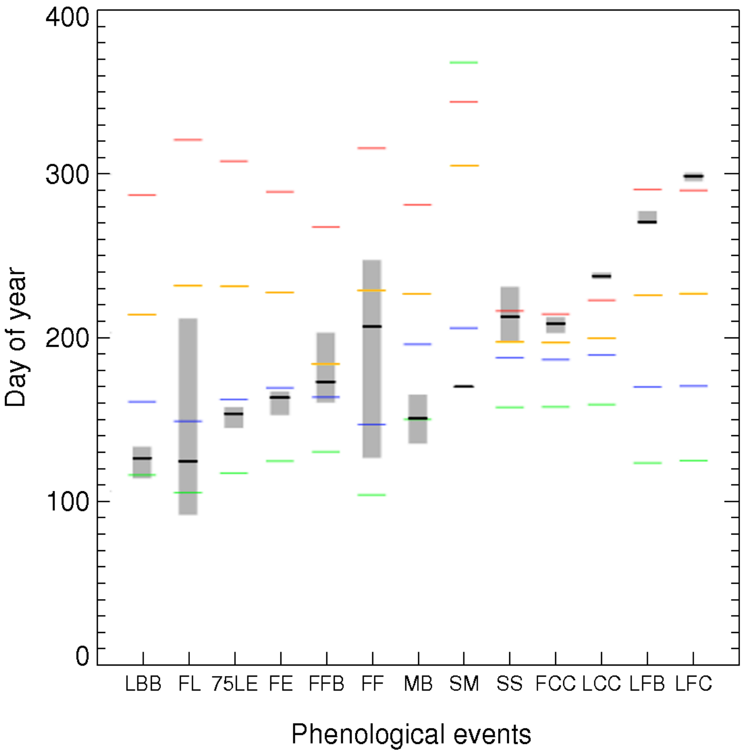


Figure 2.3: Comparison of ground and MODIS phenology. Boxplot shows ground data: shaded area is interquartile range and black line is the median; x-axis labels refer to events described in Table 2.1. Colored horizontal lines show median MODIS phenology events: green, greenup onset; blue, maturity onset; yellow, senescence onset; red, dormancy onset. As the ground phenology events transition from early season on the left to late season on the right, the MODIS phenology event closest to the median ground event also shifts toward later stages (green to red).

Table 2.2: The correlation coefficient between ground phenological events and corresponding MODIS phenology products

Phenophases	Events	N	Greenup	Maturity	Senescence	Dormancy
Greenup	Leaf bud burst	459	0.74**	0.42		
	First leaf	1287	-0.07	0.07		
	75% leaf expanding	49	0.66**	0.72**		
	Full expansion	10	-0.37	0.15		
	First flowers bud	100	-0.41	-0.30		
	First flower	2829	-0.26	0.01		
	Middle bloom	1513	0.14	-0.09		
Seeding	Seed maturity	2				
	Seed shedding	150		0.17	0.19	
Senescence	First leaf color change	445			0.00	-0.03
	Leaf fall begin	13			0.47**	-0.20
	Last leaf color change	143			0.10	-0.06
	Leaf fall complete	11			-0.18	0.71**

** Correlation passed 0.05 significant test.

Although the blooming of specific species is commonly observed in the spatial extensive network, as I addressed before, flowering is often not coincident with greenup, especially for species extensive network. The relationship between ground phenological events and onset of corresponding MODIS phenophases is shown in Table 2.2. Two greenup events “Leaf bud burst” and “75% leaf expanding” have significant correlations to the onset of greenup from MODIS, while the events related to flowering have no such significant correlation to MODIS indexes.

"Leaf fall begin" is related to onset of senescence, as well as "Leaf fall complete" to the onset of dormancy. I present a more detailed comparison between phenological events and LSP in the following sections.

Table 2.3: The comparison of MODIS phenophase onset time with all corresponding ground phenology, and the same comparison with the earliest events at each location.

MODIS	**Greenup onset**		**Maturity onset**		**Senescence onset**		**Dormancy onset**	
Events[a]	LBB, FL, 75LE, FE, FLB, FF, MB		SM, SS		FCC, LFB, LCC, LFC		FCC, LFB, LCC, LFC	
Selection	All	Early	All	Early	All	Early	All	Late
n	6247	773	136	6	408	15	144	15
r	-0.09	0.52**	0.56**	0.99	0.37	0.28	0.63**	0.92**
Bias (day)	41.14	3.11	25.94	0	12.08	17.07	14.69	3.27
RMSE (day)	66.40	23.16	33.06	12.6	19.03	49.47	25.98	14.22

[a] The abbreviation of phenology events: LBB (Leaf bud burst); FL (First leaf); 75LE (75% leaf expanding); FE (Full expansion); FLB (First flower bud); FF (First flower); MB (Middle blooming); SM (Seed maturity); SS (Seed shedding); FCC (First leaf color change); LFB (Leaf fall begin); LCC (Last leaf color change); LFC (Leaf fall complete).

** Correlation passed 0.005 significant tests.

2.3.1 Comparison between MCD12Q2 greenup and ground greenup events

Spring phenological events are classified as crucial events in all ground measurements. There is, however, no significant relationship between the MODIS greenup onset date and all spring phenological events when combined (r=-0.09, Table 2.3). Conversely, the earliest greenup events chosen from each year-location cluster are related to the MODIS greenup onset date with a statistically significant correlation (r=0.51, Table 2.3). The earliest phenology events for year-location clusters include "Leaf bud burst," "First leaf," "First flowering bud," "First flower," and

"Middle blooming." The RMA analyses of earliest phenology events and MODIS onset of greenup provided a regression slope close to 1:1 line by considering errors from both of them (Figure 2.4).

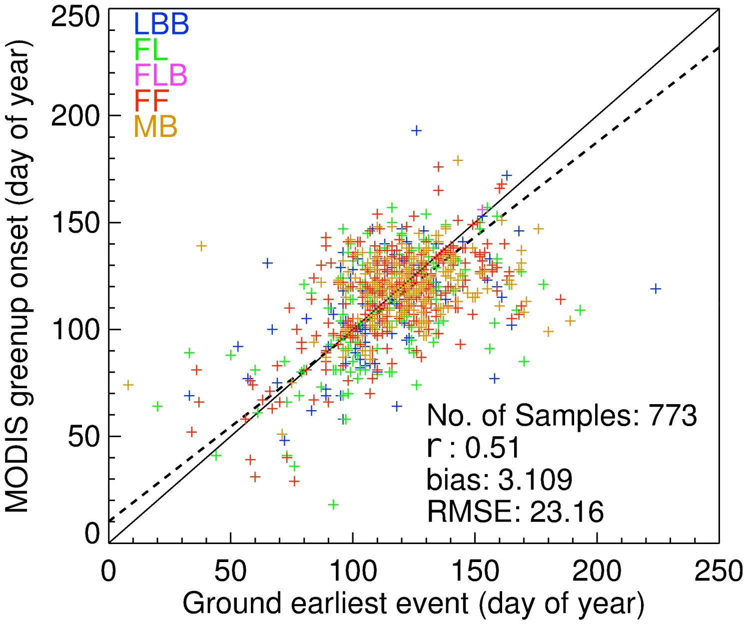


Figure 2.4: The earliest greenup ground phenological events at all sites, and corresponding MODIS onset of greenup. Dotted line is RMA regression: GreenupMODIS = 10.13+0.89* Greenupground, $p < 0.01$. The abbreviation of phenology events: LBB (Leaf bud burst); FL (First leaf); FLB (First flower bud); FF (First flower); MB (Middle blooming)

I employed a separate exercise of phenology events and MODIS changes along latitude. Calculated by latitudinal bands, the average of all the MODIS greenup pixels agreed well with the corresponding earliest ground phenology event (Figure 2.5); the MODIS-based estimate showing more variance relative to the ground-based estimates. To represent the phenology of a whole plant community, I plotted the year-location clusters (c.f. Section 2.2.1) with more than 30 phenological events across taxa (Figure 2.6). Although the frequency distributions of all spring phenological events in different sites varied in amplitude and dispersion, the corresponding MODIS onset of greenup is close to the beginning of the spring event frequency distribution for most of the year-location clusters, supporting the results shown in table 2.3. This implies that the MODIS-based onset of greenup matches the earliest ground phenological events in the community. In addition, the MODIS onset of greenup was more closely related to the earliest ground-based greenup events for sites with more than five species in the record, where the extensive observed species have a better representation of local vegetation (r=0.61, Figure 2.7).

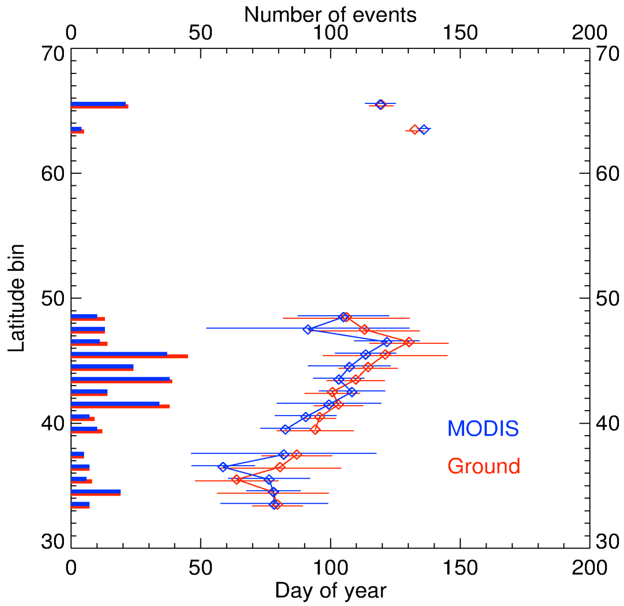


Figure 2.5: Average earliest ground phenological events (red line and diamond) and the corresponding average MODIS onset of greenup (blue line and diamond) calculated by latitude. Standard deviations are shown as horizontal lines. Bars on the left show the number of available samples (red, ground; blue, MODIS) in each latitude bin.

Lilac is often used as an indicator species for phenological observation, and it is planted and observed in some large networks and even in international phenology networks. The “First flower” of lilac is a relatively early event in temperate region, and thus might be considered as an indicator of the start of growing season for most communities in North America. Yet the species tends to keep its own development rhythm, which may not be the same as the dynamics as observed from its respective local community. The “First flower” of Lilac may occur at a relatively late period in some regions, and may not reflect the onset of a community greenup. Since satellite sensors gather integrated information from the land surface, it is also questionable whether the MODIS phenology is related to the phenology of selected single species. I compared the MODIS onset of greenup with “First flower” of common lilac (different genetic stock) and cloned lilac (same genetic characteristics) from different observational networks. Results show that selected species of lilacs are comparable to MODIS-based greenup estimates from day of year 60 to 150. But MODIS underestimates greenup by about 20 days for lilac flowering after day of year 150, and the “First flower” of lilacs occurs much later when the MODIS onset of greenup is below 60 day of year (Figure 2.8).

The comparative results for greenup show that the MODIS phenology product captures the earliest events of plant community growth. Of all the plant phenological stages, the greenup events on leaves are most consistently related to the MODIS-estimated greenup. The earliest blooming of plants also has a high correlation with MODIS-estimated greenup. Due to the previously discussed uncertainties, the blooming dynamics of a whole community may not be a good candidate for representing greenup. Lilac early blooming is also significantly correlated with MODIS-based greenup but it is still questionable whether a singular species event can represent a plant community phenology across large regions.

2.3.2 Comparison for maturity, senescence, and dormancy

Four events are recorded during the senescence phenophases in ground observations: “First leaf color change,” “Leaf fall begin,” “Last leaf color change,” and the “Completion of leaf fall.” In my study, I compared the MODIS-based senescence onset timing with two relatively early plant senescence events, “First leaf color change” and “Leaf fall begin.” Neither of the combination of plant senescence events or single events was correlated to MODIS-based senescence onset (Figure 2.9a). The data points of the “First leaf color change” against the

MODIS-based senescence onset are around 1:1 line, while the “Leaf fall begin” was independent from the senescence onset. The onset of dormancy from MODIS represents the “Last leaf color change” and “Leaf fall complete” (RMSE=25.98 day for all events, Figure 2.9b, Table 2.3), and the last events as recorded in the “leaf fall complete” and “last leaf color change” resemble MODIS-derived dormancy onset (RMSE=14.22 day, Table 2.3). The dates of all seeding are related to the MODIS-based maturity onset (r=0.56, $p<0.05$, Table 2.3), in spite of only a few seeding records at three locations in the dataset (Figure 2.9c).

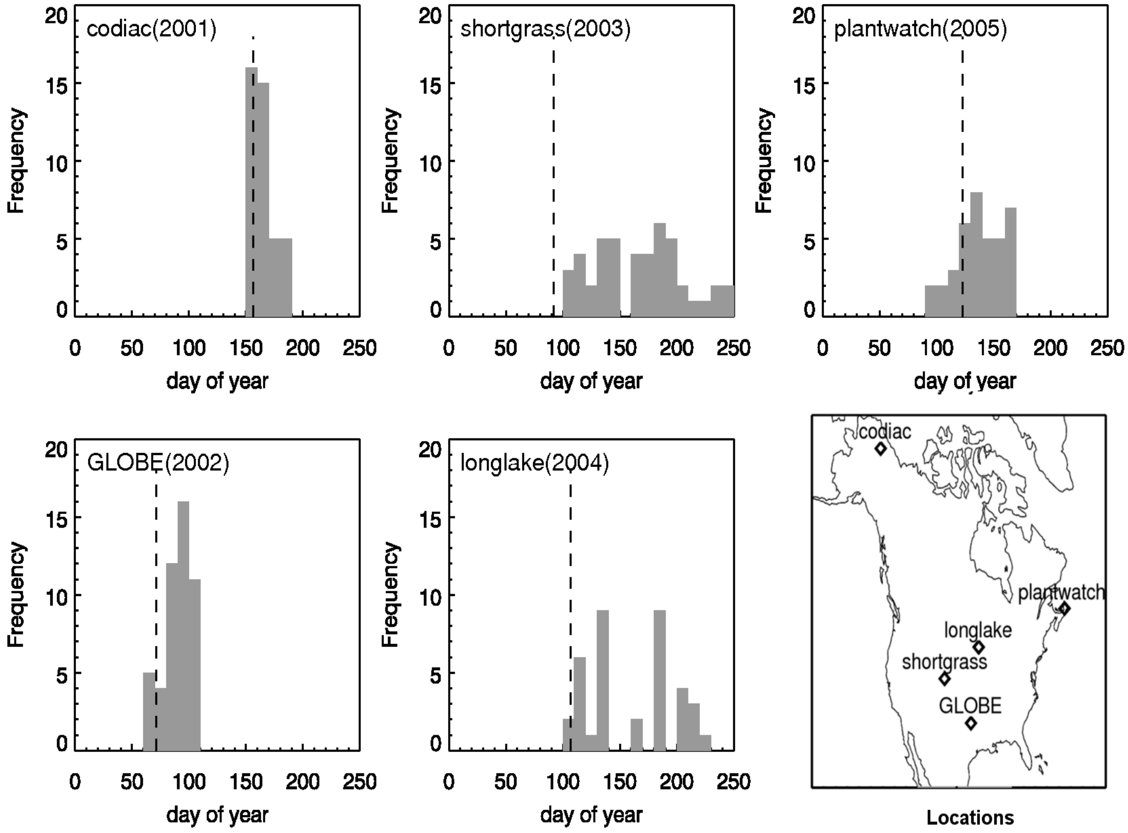


Figure 2.6: The frequency of spring phenological events (histogram bars at 10-day bins) and MODIS onset date of greenup (vertical dotted line) in five sites with more than 30 species. Locations of five sites are showed in lower right plot.

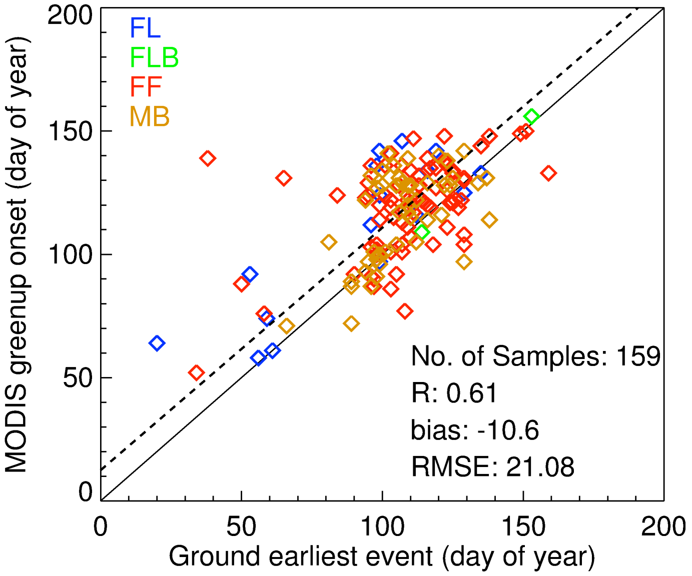


Figure 2.7: The earliest greenup ground phenological events on the sites observed more than five species, and corresponding MODIS onset of greenup. RMA regression showed in dotted line: $Greenup_{MODIS} = 12.45 + 0.98 * Greenup_{ground}$, $P<0.01$.

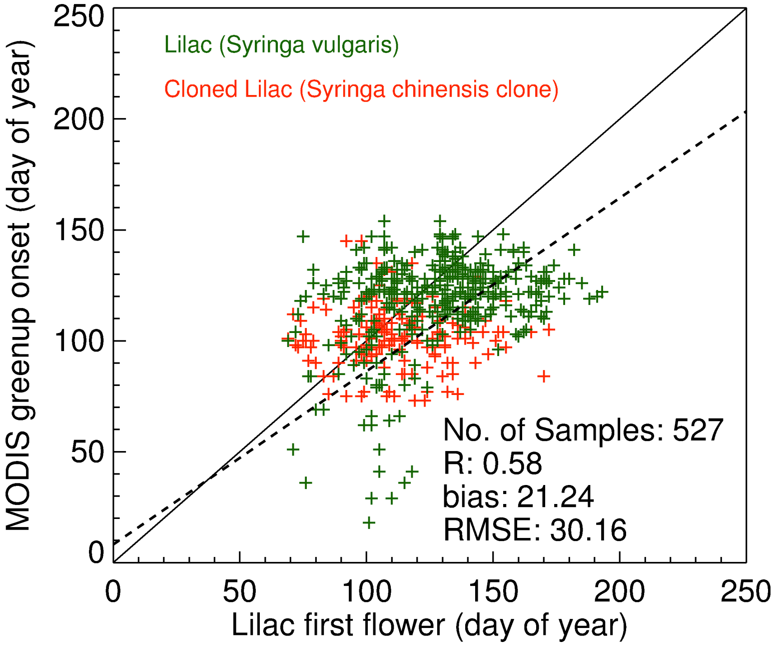


Figure 2.8: Comparison between common lilac and cloned lilac spring phenological events and onset of greenup. The RMA regression for all lilac events and MODIS greenup onset is showed in dot line: $Greenup_{MODIS} = 8.15 + 0.78 * Greenup_{ground}$, $p<0.05$.

Overall, there are relationships from some ground observed phenology events during maturity, senescence, and dormancy to the corresponding MODIS LSP metrics, for example, “First leaf color change” to onset of senescence, “Last leaf color change” to onset of dormancy, and “Developing fruit”/ “Seed maturity” to onset of maturity. However the robustness of the relationships is difficult to assess because the measurements are spatially sparse and temporally undersampled relative to the extensive measurements that exist for spring phenological events.

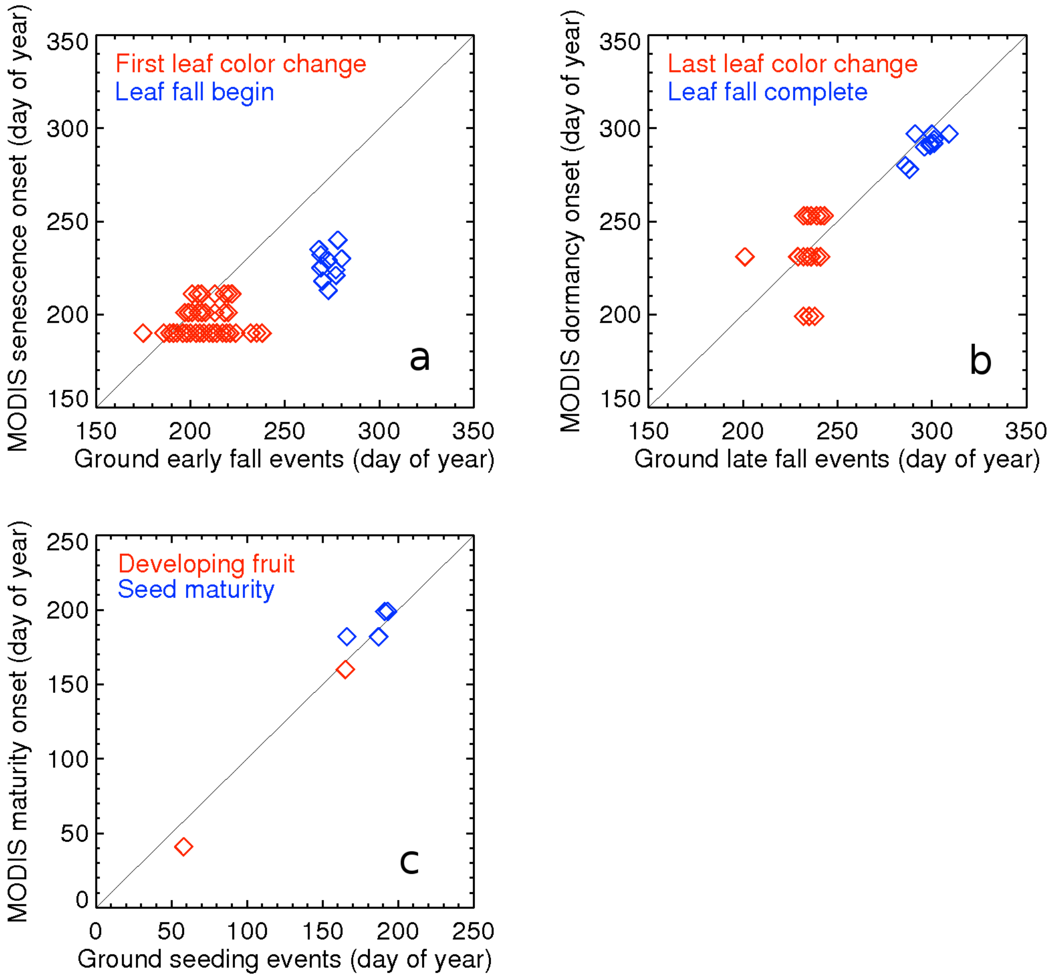


Figure 2.9: The comparison of ground measurements and MODIS indices during maturity, senescence, and dormancy. (a) MODIS senescence onset vs. ground early fall events; (b) MODIS dormancy onset vs. ground late fall events; (c) MODIS maturity onset vs. seeding events.

2.3.3 Intercomparison of greenup among MODIS, AVHRR and MERIS

The standard deviations of MODIS and AVHRR show similar spatial patterns (Figure 2.10a&b). The interannual variability of these two datasets is generally less than 10 days in the east of the CONUS, northwest Canada, and in Alaska. The standard deviation of AVHRR greenup, however, is conspicuously larger than 15 days in the north of New England compared to the low values as seen in the MODIS greenup. In the west of the US both datasets have standard deviations more than 10 days, and AVHRR is much lower in most of the eastern US (less than 5 days).

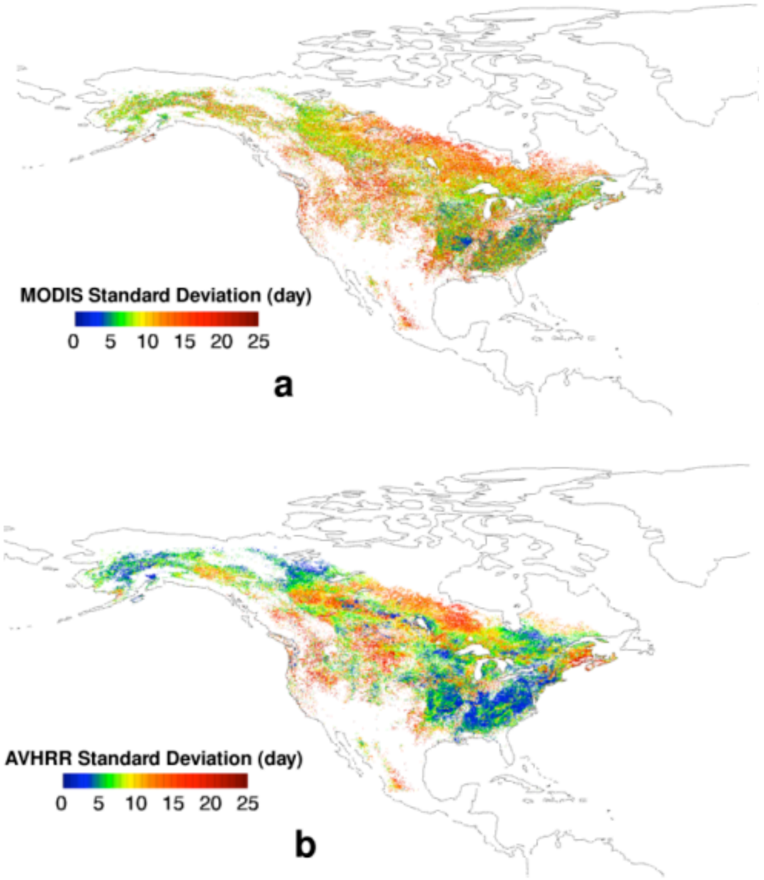


Figure 2.10: The standard deviation of MODIS greenup onset (a) and AVHRR greenup (b) from 2001 to 2006.

Due to MTCI data availability, I present MTCI comparisons for CONUS over 2003 to 2006. The variation in average timing of onset of greenup by latitude for MODIS and MTCI follows each other more closely than with does AVHRR, which is often delayed by 15 days from 31°N to 45° N (Figure 2.11). In low latitude regions below 30°N, the onset of greenup from AVHRR is often 20 days later than MODIS and MTCI values, and the differences between MODIS and MTCI are also higher. However, the AVHRR product precedes the MODIS greenup dates by about five days from 50°N to 70° N. I calculated the interannual average onset of greenup from these three sensors for all pixels in the North America from 2003 to 2006, and the correlation coefficients were calculated to intercompare each of the sensors (Figure 2.12). The correlation between MODIS and MTCI was observed to be the largest while the AVHRR and MTCI displayed a lower correlation.

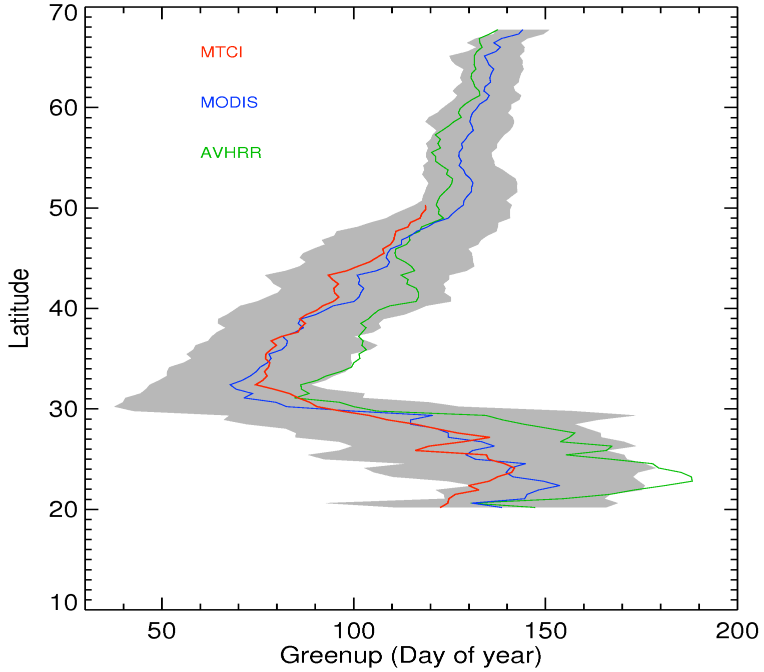


Figure 2.11: Average onset of greenup time from 2003 to 2006 based on AVHRR, MODIS, and MTCI calculated by latitude in each half degree latitude bin. The shaded area is the MODIS standard derivation.

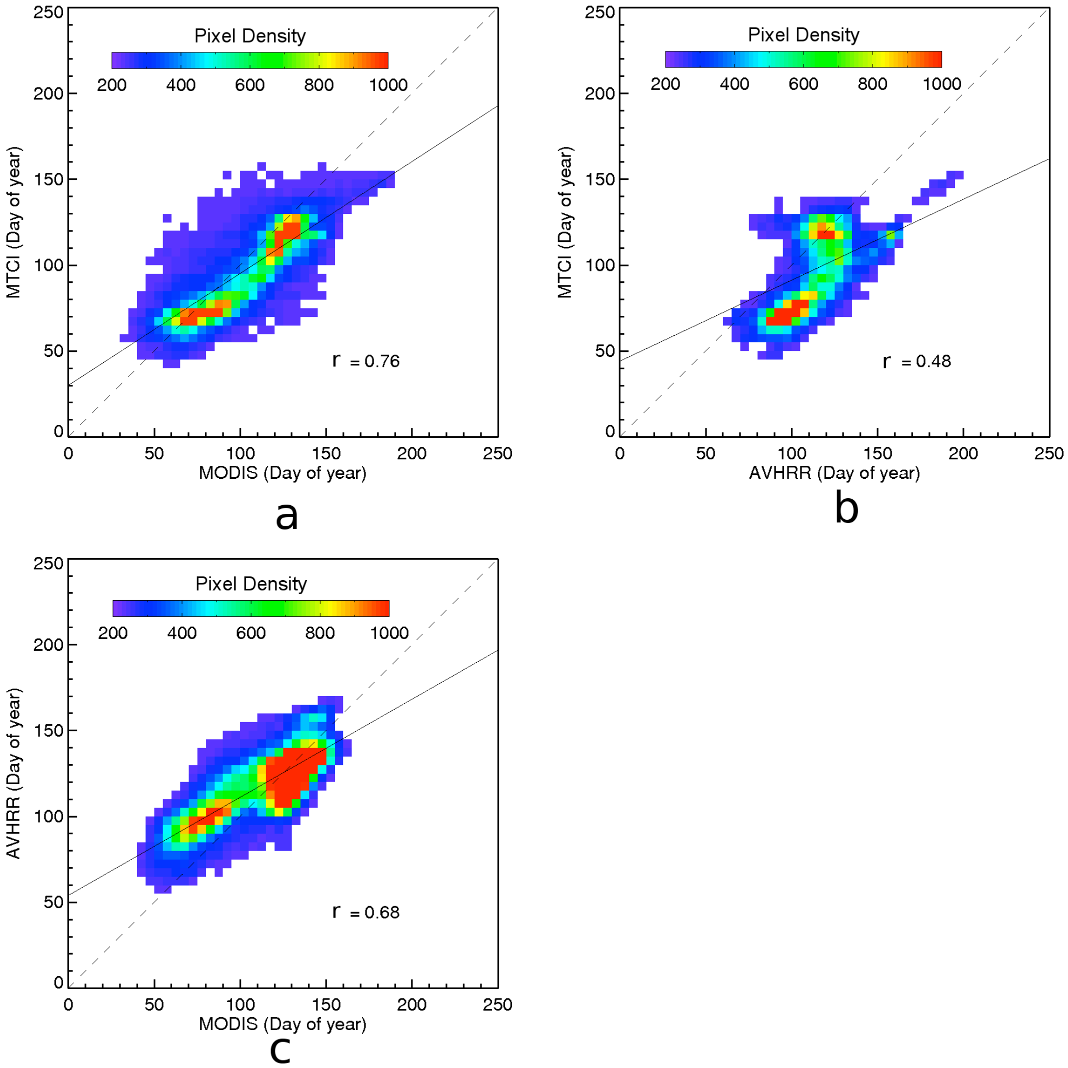


Figure 2.12: Plots of average onset of greenup in CONUS during 2003 to 2006 from different sensors and their correlation coefficient (a) MODIS and MTCI; (b) AVHRR and MTCI; (c) MODIS and AVHRR. The color shows point density in the 2D space. Dotted line is 1:1 line; dashed line is the slope line of two data.

The MODIS - MTCI comparison has the highest correlation (Figure 2.12). This is encouraging because both products involve distinct spectral combinations in vegetation indices calculation and the phenology detection algorithm. The midpoint methods used in the AVHRR greenup product reflects the transition period of greenup between dormancy and maturity by considering low VI values due to snow and bare soil, whereas algorithms of MODIS and MTCI greenup products tend to capture the onset of greenup (beginning of increasing trend on VI curves) as their goal. Hence, AVHRR greenup indices tend to occur later than MODIS and MTCI greenup in the low and middle latitude region without snow cover in winter and spring. Additionally, AVHRR greenup is derived from an antiquated sensor, GIMMS is stored at a far coarser spatial resolution, and the sensors all use different band combinations: any of these factors could influence the observed differences. At least for the sites analyzed and at the specific scale of analysis undertaken here, this study implies that similar spatial patterns of greenup can be extracted from different sensors even with different methodologies and band combinations, although the exact greenup dates derived from these products differs in particular regions.

2.4 Conclusion

The intercomparison exercise in this paper shows that the MODIS onset of greenup is statistically correlated with the earliest plant phenological events during greenup in different landscapes over North America. Since the observed species are limited in spatially extensive observation networks, these networks may not capture the beginning of the growing season for various communities across North America. Besides the spring phenological events, ground measurements of other events during summer and fall are severely inadequate for interpreting the remotely sensed product, although the direct comparison in limited sites shows similarity between MODIS products and a few ground records for maturity and senescence stages.

An extended growing season due to senescence delay can lead to net carbon dioxide losses, due to enhanced respiration combined with reduced photosynthetic capacity during autumn warming (Piao et al., 2008). This urges for the construction of more ground observations networks for recording senescence events. Overall, the number of ground plant phenological records is still spatially and temporally inadequate to reveal large scale variability in growing season length, and thus there is a pressing need for more species-level observations available through a spatially extensive database. Promising efforts, such as the National Phenology Network (Betancourt et al., 2007), are already underway.

All LSP indices, despite being calculated from different sensors based on various algorithms, consistently reflect variation in phenology patterns along latitude. The MODIS onsets of greenup in North America is correlated to MTCI greenup at similar spatial resolution (R=0.76), whereas the AVHRR greenup shows less interannual variance than the ones derived from MODIS. Since the LSP may vary in different spatial resolutions especially for heterogeneous landscapes (Liang et al., 2011), phenological products of higher resolutions should provide more accurate and detailed estimates of vegetation seasonal dynamics. The notable lag of AVHRR greenup against onset of greenup from the MODIS and MTCI products implies a diverse definition in phenological indices. Therefore LSP indices require further clarification for definitions and closer connection with ground vegetation/plant phenology. Then LSP indexes amongst sensors might match each other after being carefully validated by ground measurement.

References

Badeck, F. W., Bondeau, A., Böttcher, K., Doktor, D., Lucht, W., Schaber, J., et al. (2004), Responses of spring phenology to climate change, *New Phytologist*, *162*(2), 295–309.

Berterretche, M., Hudak, A. T., Cohen, W. B., Maiersperger, T. K., Gower, S. T., & Dungan, J. (2005), Comparison of regression and geostatistical methods for mapping Leaf Area Index (LAI) with Landsat ETM+ data over a boreal forest, *Remote Sensing of Environment*, *96*(1), 49–61.

Betancourt, J. L., Schwartz, M. D., Breshears, D. D., Brewer, C. A., Frazer, G., Gross, J. E., et al. (2007), Evolving plans for the USA National Phenology Network, *Eos, Transactions American Geophysical Union*, *88*(19), 211.

De Beurs, K. M., & Henebry, G. M. (2004), Land surface phenology, climatic variation, and institutional change: Analyzing agricultural land cover change in Kazakhstan, *Remote Sensing of Environment*, *89*(4), 497–509.

Cesaraccio, C., Spano, D., Snyder, R. L., & Duce, P. (2004), Chilling and forcing model to predict bud-burst of crop and forest species, *Agricultural and forest meteorology*, *126*(1-2), 1–13.

Chmielewski, F. M., & Rötzer, T. (2001), Response of tree phenology to climate change across Europe, *Agricultural and Forest Meteorology*, *108*(2), 101–112.

Chuine, I., Yiou, P., Viovy, N., Seguin, B., Daux, V., & Ladurie, E. L. (2004), Historical phenology: grape ripening as a past climate indicator, *Nature*, *432*(7015), 289–290.

Dash, J., & Curran, P. J. (2004), The MERIS terrestrial chlorophyll index, *International Journal of Remote Sensing*, *25*(23), 5403–5413.

Dash, J., Jeganathan, C., & Atkinson, P. M. (2010), The use of MERIS Terrestrial Chlorophyll Index to study spatio-temporal variation in vegetation phenology over India, *Remote Sensing of Environment*, *114*(7), 1388–1402.

Euskirchen, E. S., McGuire, A. D., Kicklighter, D. W., Zhuang, Q., Clein, J. S., Dargaville, R. J., et al. (2006), Importance of recent shifts in soil thermal dynamics on growing season length, productivity, and carbon sequestration in terrestrial high□latitude ecosystems, Global Change Biology, 12(4), 731–750.

Fisher, J. I., & Mustard, J. F. (2007), Cross-scalar satellite phenology from ground, Landsat, and MODIS data, *Remote Sensing of Environment*, *109*(3), 261–273.

Forrest, J., Inouye, D. W., & Thomson, J. D. (2010), Flowering phenology in subalpine meadows: Does climate variation influence community co-flowering patterns?, *Ecology*, *91*(2), 431–440.

Ganguly, S., Friedl, M. A., Tan. B., Zhang, X., & Verma, M. (2010), Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product, *Remote Sensing of Environment*, *114*(8), 1805–1816.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002), Overview of the radiometric and biophysical performance of the MODIS vegetation indices, *Remote Sensing of Environment*, *83*(1-2), 195–213.

Hufkens, K., Friedl, M., Sonnentag, O., Braswell, B. H., Milliman, T., & Richardson, A. D. (2012), Linking near-surface and satellite remote sensing measurements of deciduous broadleaf forest phenology, *Remote Sensing of Environment*, *117*(0), 307–321.

Ishida, A., Diloksumpun, S., Ladpala, P., Staporn, D., Panuthai, S., Gamo, M., et al. (2006), Contrasting seasonal leaf habits of canopy trees between tropical dry-deciduous and evergreen forests in Thailand, *Tree Physiology*, *26*(5), 643–656.

Keeling, C. D., Chin, J. F. S., & Whorf, T. P. (1996), Increased activity of northern vegetation inferred from atmospheric CO2 measurements, *Nature*, *382*(6587), 146–149.

Kimball, J. S., McDonald, K. C., Frolking, S., & Running, S. W. (2004), Radar remote sensing of the spring thaw transition across a boreal landscape, *Remote Sensing of Environment*, *89*(2), 163–175.

Körner, C., & Basler, D. (2010), Phenology under global warming, *Science*, *327*(5972), 1461.

Liang, L., & Schwartz, M. D. (2009), Landscape phenology: an integrative approach to seasonal vegetation dynamics, *Landscape ecology*, *24*(4), 465–472.

Liang, L., Schwartz, M. D., & Fei, S. (2011), Validating satellite phenology through intensive ground observation and landscape scaling in a mixed seasonal forest, *Remote Sensing of Environment*, *115*(1), 143–157.

Morisette, J. T., Richardson, A. D., Knapp, A. K., Fisher, J. I., Graham, E. A., Abatzoglou, J., et al. (2008), Tracking the rhythm of the seasons in the face of global change: phenological research in the 21st century, *Frontiers in Ecology and the Environment*, *7*(5), 253–260.

Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G., & Nemani, R. R. (1997), Increased plant growth in the northern high latitudes from 1981 to 1991, *Nature*, *386*(6626), 698–702.

Peñuelas, J., & Filella, I. (2009), Phenology feedbacks on climate change, *Science*, *324*(5929), 887.

Piao, S., Ciais, P., Friedlingstein, P., Peylin, P., Reichstein, M., Luyssaert, S., et al. (2008), Net carbon dioxide losses of northern ecosystems in response to autumn warming, *Nature*, *451*(7174), 49–52.

Reed, B. C., Brown, J. F., VanderZee, D., Loveland, T. R., Merchant, J. W., & Ohlen, D. O. (1994), Measuring phenological variability from satellite imagery, *Journal of Vegetation Science*, *5*(5), 703–714.

Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., & Ohno, H. (2005), A crop phenology detection method using time-series MODIS data, *Remote sensing of environment*, *96*(3), 366–374.

Schwartz, M. D. (1994), Monitoring global change with phenology: the case of the spring green wave, *International journal of biometeorology*, *38*(1), 18–22.

Schwartz, M. D. (2003), *Phenology: An Integrative Environmental Science*, Springer-Verlag, Berlin.

Schwartz, M. D., & Hanes, J. M. (2010), Intercomparing multiple measures of the onset of spring in eastern North America, *International Journal of Climatology*, *30*(11), 1614–1626.

Schwartz, M. D., & Reiter, B. E. (2000), Changes in north American spring, *International Journal of Climatology*, *20*(8), 929–932.

Schwartz, M. D., Ahas, R., & Aasa, A. (2006), Onset of spring starting earlier across the Northern Hemisphere, *Global Change Biology*, *12*(2), 343–351.

Sokal, R. R., & Rohlf, F. J. (1995), *Biometry: The Principles and Practice of Statistics in Biological Research*, WH Freeman, New York. 892

Thompson, D. R., & Wehmanen, O. A. (1979), Using Landsat digital data to detect moisture stress, *Photogrammetric Engineering and Remote Sensing. 45:*201-207.

Tucker, C. J., Pinzon, J. E., Brown, M. E., Slayback, D. A., Pak, E. W., Mahoney, R., et al. (2005), An extended AVHRR 8□km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data, *International Journal of Remote Sensing*, 26(20), 4485–4498.

White, M. A., Thornton, P. E., & Running, S. W. (1997), A continental phenology model for monitoring vegetation responses to interannual climatic variability, *Global Biogeochemical Cycles*, *11*(2), 217–234.

White, M. A., de Beurs, K. M., Didan, K., Inouye, D. W., Richardson, A. D., Jensen, O. P., et al. (2009), Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006, *Global Change Biology*, *15*(10), 2335–2359.

Wilson, K. B., Baldocchi, D. D., & Hanson, P. J. (2000), Spatial and seasonal variability of photosynthetic parameters and their relationship to leaf nitrogen in a deciduous forest, *Tree Physiology*, *20*(9), 565.

Xu, L., Baldocchi, D. D., & Tang J. (2004), How soil moisture, rain pulses, and growth alter the response of ecosystem respiration to temperature, *Global Biogeochemical Cycles*, *18*(4), GB4002.

Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C., Gao, F., et al. (2004), Monitoring vegetation phenology using MODIS, *Remote Sensing of Environment*, *84*(3), 471–475.

Zhang, X., Friedl, M. A., & Schaaf, C. B. (2006), Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements, *Journal of Geophysical Research (Biogeosciences)*, *111*(G10), 04017.

Zhang, X., Tarpley, D., & Sullivan, J. T. (2007), Diverse responses of vegetation phenology to a warming climate, *Geophysical Research Letters*, 34(19), L19405.

CHAPTER 3

A SIMPLE PARAMETRIC ESTIMATION OF LIVE FOREST ABOVEGROUND BIOMASS IN CALIFORNIA USING SATELLITE DERIVED METRICS OF CANOPY HEIGHT AND LEAF AREA INDEX

Abstract

Accurate characterization of variability and trends in forests biomass at local to national scales is required for accounting of global carbon sources and sinks, and monitoring their evolution. Here I present a new remote sensing based approach for mapping live forest aboveground biomass (AGB), independent of the field inventory data needed in many other approaches. The model is based on a simple parametric model that combines high-resolution estimates of Leaf Area Index derived from the Thematic Mapper sensor onboard the Landsat satellite and canopy maximum height from the Geoscience Laser Altimeter System (GLAS) sensor onboard ICESat, the Ice, Cloud, and land Elevation Satellite. I tested my model over the forested areas of California. spanning a broad range of climatic and land-use conditions, and show that the AGB estimates are comparable to inventory estimates. The method is particularly useful for areas with underrepresented or outdated field measurements, and may help to minimize inconsistencies in AGB estimates caused by methodological variations in field inventories across nations. My approach offers a high-resolution approach for forest carbon monitoring with the potential of global applicability.

3.1 Introduction and Motivation

Quantifying and preserving the carbon stock in forests underpins sequestration and mitigation strategies, such as the program to Reduce Emissions from Deforestation and Degradation (REDD) (UNFCCC; Canadell et al., 2007; IPCC, 2007; Miles and Kapos, 2008), and promotes carbon credits to offset greenhouse gas emissions from other sources. Forest carbon stocks change over time and can be monitored through periodic mapping of the total forest biomass (Gurney and Raymond, 2008; Saatchi et al., 2011; Baccini et al., 2012). National inventories, e.g. the Forest Inventory and Analysis (FIA) Program of the U.S. Forest Service, provide the most detailed and accurate field observations of forest biomass and disturbance history and thus are a key information source in many carbon studies (Keith et al., 2010). Yet the utility of such inventories for global-scale studies is limited because of national differences in measurement strategies and, in the case of remote forests such as those of the boreal and tropical regions, sparse sampling and out of date observations (Houghton, 2005).

Recent efforts relying on remote sensing allow mapping the distributions of the aboveground portion of total forest biomass (AGB) over wide geographical extents, overcoming problems of field data underrepresentation (Lu et al., 2006; Gibbs et al., 2007). Canopy reflectance measured by passive optical sensors and radar backscatter may be correlated with field observed AGB (Gibbs et al., 2007; Saatchi et al., 2007). Precision, however, is limited for AGB densities greater than 300-400 Mg/Ha, as these reflective measures show an asymptotic behavior with canopy closure (Saatchi et al., 2007). Canopy 3-D modeling using space-borne Lidar and Interferometric synthetic aperture radar (InSAR) have improved the accuracy of the estimates of canopy vertical structure, with Lidar tending to outperform InSAR (Lefsky et al., 1999; Drake et al., 2002; Asner et al., 2010). However, the sparse availability of Lidar data requires the fusion of multiple sources of data for large area mapping of AGB. Approaches include the combination of forest height derived from the Geoscience Laser Altimeter System (GLAS, onboard ICESat, the Ice, Cloud, and land Elevation Satellite Harding & Carabajal, 2005), reflected energy (e.g. from the Moderate Resolution Imaging Spectroradiometer - MODIS and Phased Array type L-band Synthetic Aperture Radar - PALSAR) and ground data, all processed with predictive models (Baccini et al., 2008; Boudreau et al., 2008; Goetz et al., 2009; Saatchi et al., 2011; Baccini et al., 2012). The field-data driven model estimates of AGB have large uncertainties in regions where inventory data are absent and the inferences on under-sampled areas are usually based on unified allometric models derived from a few representative sites. To circumvent these problems, in this study I develop a methodology to estimate AGB based on the following steps:

(a) Produce a robust continuous height map using a simple parametric model linking Landsat-derived leaf area index (LAI) to forest height from GLAS;

(b) Estimate AGB using a height-to-biomass functional relationship; and

(c) Compare my AGB estimates to inventory data and other satellite-derived maps of AGB available at different spatial resolutions.

In the following sections I describe the region of study, data and methods used to estimate AGB and approaches to inter-compare AGB estimates with inventory based plot data. Detailed information on data and methods is provided in the auxiliary material and references to sections in the auxiliary are provided here for legibility.

3.2 Data and Methods

The study area for prototyping my methodology is the state of California, which encompasses various eco-climatic zones with large variations in AGB. The total forested land in California is approximately 9.54 million hectares (Section 3.5.1.2.1 and Figure 3.4a), with evergreen being the most common forest type. I define the structural attributes of the forested pixels through (1) the LAI derived from the Landsat 5 TM surface reflectance data in the Red and Near-Infrared (NIR) spectral bands and (2) the potential canopy height obtained from the available ICESat GLAS data. At the scale of a Landsat pixel, I assume that the LAI represents an average of the vertically integrated leaf density of the tree stands and correlates with the average canopy height within homogeneous climate zones and stand compositions. A list of all the data used in this study is also detailed in Table 3.4.

I used the available cloud-free (cloud cover less than 5%) Landsat 5 TM scenes from the USGS Global Land Survey (GLS) 2005 dataset covering the state of California. I identified images with low residual noise and nearly matching acquisition dates during the growing season (May to October) for the years from 2004 to 2006 to minimize phenological changes across the different scenes. The images were pre-processed using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) for atmospheric correction and calibration to compute surface spectral reflectances (Section 3.5.1.1.1). The surface reflectances, combined with National Land Cover Database land cover data, were used to retrieve LAI for the forested pixels using a radiative transfer based inversion approach (Section 3.5.1.1.2; Ganguly et al., 2008; Ganguly et al., 2012).

I analyzed the 2003 to 2007 GLAS GLA14 data to assemble the best possible imagery (Section 3.5.1.1.3). GLAS emits a 1064 nm laser pulse at an elliptical ground footprint of ~64 m diameter (Abshire et al., 2005), and the echo waveform of the laser pulse at 0.15 m vertical sampling resolution contains information related to canopy structure, the land surface, snow coverage, and so on. I derived maximum canopy height from each of the GLAS waveforms (Figure 3.1a) and assessed the quality of the height values using accompanying flags as well as ancillary data from the National Elevation Data (NED) and Landsat (Section 3.5.2.1). My strict filtering approach resulted in 8193 GLAS shots that are spatially representative of the forests of California. A density scatter plot between the GLAS obtained maximum height (Figure 3.1a) and the Landsat LAI (Figure 3.5a) nearest to the center coordinates of the GLAS shots shows a moderate but highly significant linear correlation (r=0.521, p value = 2.2x10-6) (Section 3.5.3). A similar exercise with the Normalized Difference Vegetation Index (NDVI) shows a weaker correlation because of saturation at high canopy height values. LAI has comparatively less saturation (Figure 3.5b).

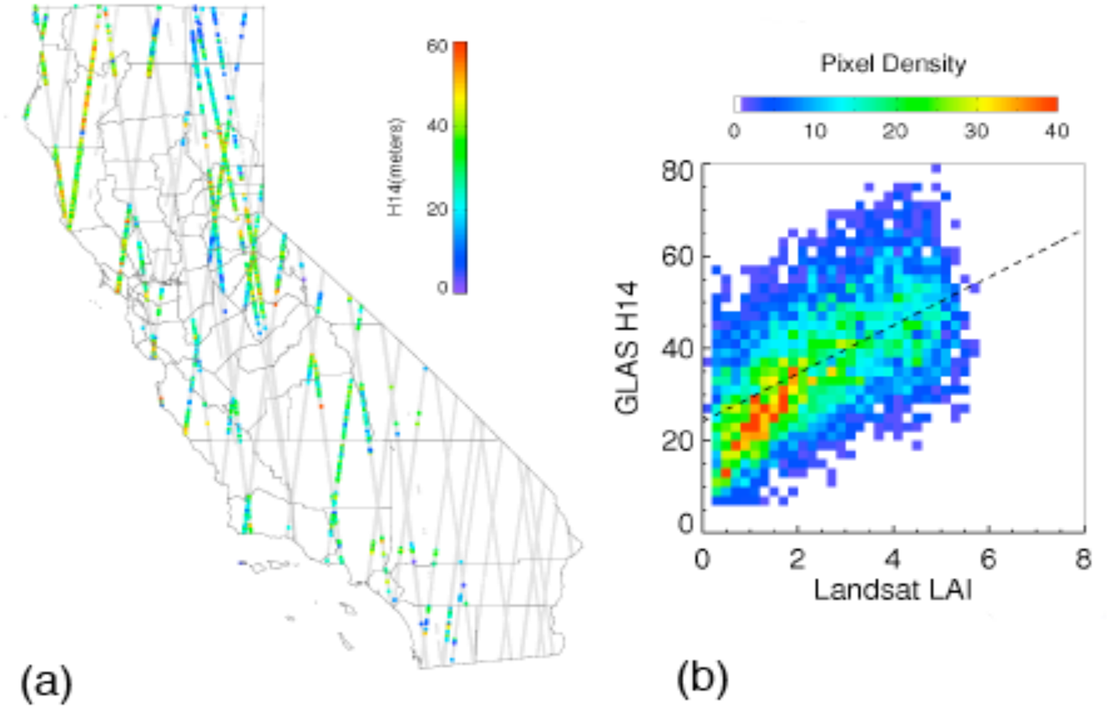


Figure 3.1: Estimating forest height in California. (a) Geographic distribution of GLAS shots and retrieved canopy height values for the forested region of California. The color bar shows retrieved maximum canopy height values (H14) calculated as the difference between the signal start (zs) and ground peak elevation (zg) as estimated from a GLAS waveform. The gray color in the shot profiles shows filtered GLAS data that are invalid (see Section 3.5.2.1) (b) Density scatter plot between GLAS maximum canopy height (H14) and Landsat LAI nearest to the GLAS center locations. The total number of sample points is 8196. The fitted model is "H14 = 24.097+5.22*LAI" and the RMSE is 12.327 with a correlation of 0.521.

Capitalizing on the statistical correlation between LAI and maximum height from GLAS, I use a linear model to derive height values for the remaining 30 m forested pixels with a valid LAI (Figure 3.2a). The resulting continuous maximum height fields are then converted to a live AGB estimate (referred to as "ARC", for Ames Research Center, from this point onwards; Figure 3.2b) using a height-to-biomass functional relationship derived from the available FIA plot data for California (Section 3.5.1.3 & 3.5.4; Figure 3.8). I assess the uncertainty in the final AGB estimates through a simple error propagation model (Section 3.5.7) that analyzes the effect of errors in the input observation data and the height prediction model.

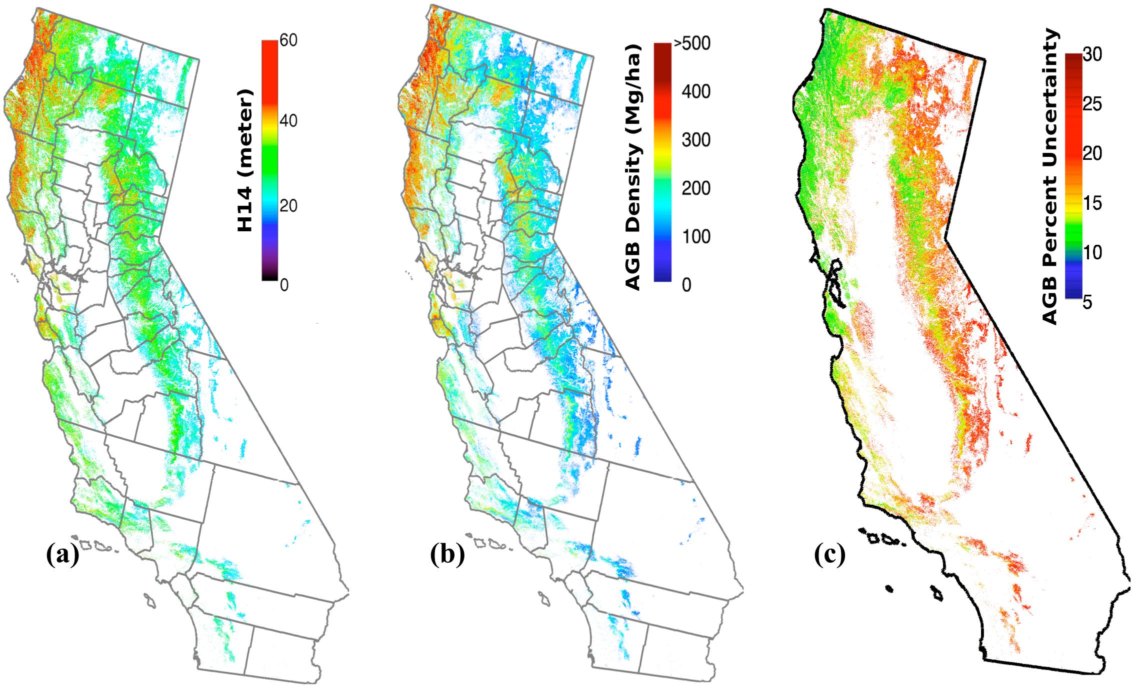


Figure 3.2: Height and biomass for California forests. (a) Continuous height derived using the model fit relationship obtained from Figure 3.1b. (b) AGB estimates using a height-biomass model derived from FIA plots. All 2205 FIA plots for California (Figure 3.7) were acquired for 2004 to 2006. A relationship between FIA estimates of height and aboveground biomass was derived for all the plots. The model relationship is applied to each forested pixel with a valid height value. (c) Percent uncertainty in AGB estimates. The uncertainty is calculated by propagating a uniform bias of LAI = 0.5 and by choosing a linear fit model for estimating height from LAI (refer to Section 3.5.7 for more details).

It is imperative to compare the spatial estimates AGB with available field data and inter-compare with other existing satellite-derived biomass maps (identified as “USFS” and “NBCD”; Refer to Table 3.4 and Section 3.5.1.4 for detailed description). To this end, I stress that plot- and pixel-level AGB estimates are not directly comparable (Section 3.5.5; Baccini et al., 2011). A fundamental step therefore is to scale the FIA plot-level data and the remotely sensed data to an aggregated or stratified scale. To validate the satellite-derived maps with FIA, I aggregate the AGB to county and sub-ecoregion total biomass (Section S6; Section 3.5.1.3). I also compare the estimates by grouping the sub-ecoregions in three biomass classes (low, medium and high) for the dominant forested ecosystem provinces (Table 3.5).

3.3 Results and Discussion

Figure 3.3a shows the frequency distributions of AGB from the satellite-derived maps. The mean behavior and the dispersion in the AGB estimates show notable disagreements when compared at their respective spatial scales. Figure 3.9 tests the hypothesis that scaling ABG across different spatial resolutions is an incorrect step and that maps of AGB at different spatial resolution cannot be directly compared. Figure 3.9 shows that ARC AGB values decrease when the native 30-m pixel estimates are re-sampled to coarser resolutions (~10.6% decrease at the USFS equivalent spatial resolution; Section 3.5.5), while the forest cover remains nearly constant (~0.2% relative change due to resampling). When comparing the FIA total biomass estimates with the satellite-derived maps for all the counties and sub-ecoregions (Figure 3.3b,c and Table 3.1), ARC shows the least absolute bias (~0.41 M ton for county total and ~0.35 M ton for sub-ecoregions) and error from measured values (also see Figure 3.10 for spatial patterns of total biomass estimated at the sub-ecoregion level). The statistics reported in Table 3.1 show the correlation coefficient, bias, and RMSE between the ABG derived from satellite and from FIA summarized by both county and sub-ecoregion for the whole study area. My error propagation model results (Section 3.5.6) suggest that for a uniform bias of 0.5 LAI units and using a linear height prediction model, the uncertainty of AGB at the pixel level ranges from ±5% (for high

AGB) to ±30% (for low AGB) (Figure 3.2c), but is constrained to less than ±20% for the entire range of total AGB at the typical county (Figure 3.12c) and the sub-ecoregion level.

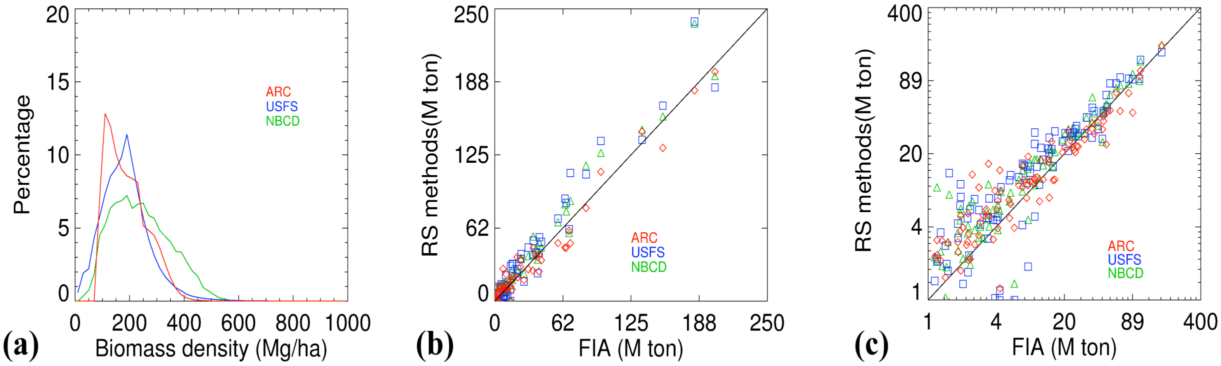


Figure 3.3: Evaluation of AGB estimates. (a) Frequency distribution of pixel-level AGB estimates from ARC, USFS and NBCD. All the maps have been reprojected to an Albers Equal Area projection scheme at a spatial grid resolution of 30-m. (b) Comparison of satellite-derived aggregated estimates of total AGB and FIA total AGB for counties. The county total AGB from the satellite-derived maps is calculated as, where Area is the pixel area, BD is the Biomass Density, and Pn is the number of pixels for the respective county. (c) Similar comparison as (b) but for sub-ecoregions. The AGB values in the plot are presented in a log scale to elucidate lower dynamic range of AGB value. The statistical parameters for both county and sub-ecoregion comparisons with FIA are shown in Table 3.1 & 3.2.

Table 3.1: Statistics for county and sub-ecoregion total biomass estimated from FIA and satellite derived estimates. All the counties and sub-ecoregions for the state of California with a valid total biomass value were included in the analysis.

Metrics[a] (w.r.t. FIA)	**ARC**	**NBCD**	**USFS**
County			
R^2	0.963	0.965	0.944
Bias	-0.413	5.209	5.360
RMSE	8.630	11.597	14.174
Sub-ecoregion			
R^2	0.940	0.966	0.938
Bias	-0.346	5.710	5.99
RMSE	8.379	9.113	11.304

[a] The Abbreviation of metriecs is: ACR for this study; NBCD for National Biomass and Carbon Database; USFS for the results from Blackard et al (2008)

Table 3.2: Sub-ecoregion total AGB difference between remote sensing methods and FIA as obtained in Figure 3.8 for three broad class ranges (Figure 3.3c). The ACCURACY relates to the mean value of relative total biomass difference at sub-ecoregion level ($\frac{1}{N}\sum_{i=1}^{N}\frac{(RS_i - FIA_i)}{FIA_i}\times 100$), while the BIAS represents the mean value of total biomass difference in M tons expressed ($\frac{1}{N}\sum_{i=1}^{N}(RS_i - FIA_i)$). Here N refers to the total number of sub-ecoregions in the respective total biomass range; RS refers to the remotely sensed estimates from ARC/NBCD/USFS. The RMSE is also reported in addition to the bias and accuracy.

Total Biomass Range (M Ton; also see Figure 3.3c)	**Number of Sub-ecoregions**	**ARC**	**NBCD**	**USFS**
BIAS				
<4.5	50	1.372	0.941	0.876
4.5 - 20	33	1.391	2.647	3.800
>20	27	-5.171	5.976	5.952
RMSE				
<4.5	50	2.89	2.07	2.83
4.5 - 20	33	4.44	4.90	7.71
>20	27	12.2	11.62	13.54
ACCURACY (%)				
<4.5	50	190.5	296.2	83.2
4.5 - 20	33	19.2	23.6	32.6
>20	27	-8.8	15.2	14.5

Table 3.3: The percentage of sub-ecoregions from remotely sensed total biomass estimates that match the FIA estimates for relative difference thresholds obtained from Figure 3.11. The sub-ecoregions in this calculation are represented by the four ecosystem provinces (M261, 263, 261 & M262 from Table 3.5) with forest as the dominant land cover.

Error Range	**ARC**	**NBCD**	**USFS**
<10%	17.3	13.6	10.0
<20%	36.4	27.3	14.5
<30%	44.5	37.3	21.8
<40%	50.9	42.7	32.7
<50%	59.1	53.6	44.5

Table 3.2 provides the bias, error, and accuracy of the satellite-derived estimates as compared to FIA for three biomass classes. In terms of bias, most of the satellite-derived maps agree to FIA with bias less than 4 M ton for sub-ecoregions with more than 4.5 M ton of biomass, with the ARC showing the least bias (averagely 1.391 M ton). In the lower biomass category (<4.5 M ton), USFS shows the lowest bias. The RMSE of ARC shows less variability across all the maps and biomass classes. ARC has improved accuracies for classes >4.5M ton while the USFS shows a smaller relative error for the lower biomass class. Figure 3.11 shows spatial disparities between the satellite-derived maps and FIA estimates at an aggregated scale. Overall, the number of sub-ecoregions that show an agreement in total AGB with FIA is highest for ARC followed by NBCD and USFS (Table 3.3). Possible explanations for the disagreement are (a) the definition and extent of forested land within the sub-ecoregions; (b) a spatial pixel underrepresentation due to fewer FIA samples.

3.4 Conclusions

In this study I provide a method to estimate live AGB based on a simple parametric model. My model builds upon the premise that high-resolution estimates of LAI and canopy maximum height can be used to estimate AGB, and I show that at an aggregated scale such estimates are comparable to inventory estimates. I feel encouraged to re-iterate the following conclusions from this study:

(a) My simple parametric model can estimate biomass values independent of field-measured values with an acceptable uncertainty, when compared to FIA regional biomass assessment. This is particularly useful for areas where field measurements are lacking and/or sparse.

(b) My aggregated AGB estimates are comparable to inventory-based estimates and available satellite derived maps.

(c) I enumerate ways to compare AGB maps with different spatial resolution at an aggregated scale.

The results from this study are in support of the ongoing activities of NASA's National Carbon Monitoring Assessment Program (CMS), which seeks to develop data fusion strategies for AGB estimation over regions at a moderately high spatial resolution with sparse or no field data. Limitations in my approach may result from input data availability (e.g. Landsat acquisition), corruption of data due to atmospheric contamination, and saturation effects in pixel retrieved surface reflectances.

3.5 Auxiliary

This auxiliary section includes detailed information about the data used, the methods applied to determine the validity of GLAS derived height metrics, the derivation of the height-biomass allometric model, and the description of methods to study the changes in aboveground biomass (AGB) density with spatial scale and to calculate total biomass estimates at county and sub-ecoregion level. An exposition on uncertainty modeling is also presented. These details are not provided in the main text for brevity. Also included are Figures 3.4-3.12 and Tables 3.4-3.7.

3.5.1 Data

3.5.1.1 Satellite Data Sets

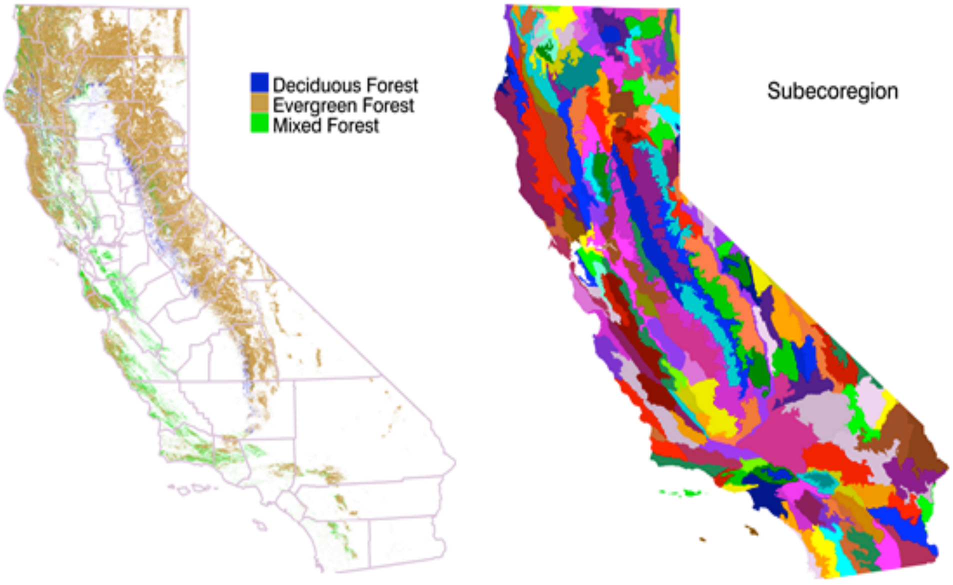
Data from a variety of satellite sensor systems are used in this study. A detailed description of the various datasets used and derived variables are provided below.

3.5.1.1.1 Landsat Surface Reflectance Data

The 2005 Global Land Survey (GLS) is a global data set of orthorectified Landsat data with core acquisition dates from 2005 to 2006. GLS 2005 consists of both Landsat 5 and gap-filled Landsat 7 imagery with a spatial resolution of 30-m (Gutman et al., 2008; WWW1). Each scene consists of a single day, cloud-free acquisition, and the time of acquisition generally falls during leaf-on conditions for the location. Data recorded in 2004 and 2007 are used to fill areas of low image quality or excessive cloud cover. In this study, we used the GLS 2005 data for the state of California (45 Landsat scenes in total). Most of the California scenes are acquired during the peak of the growing season, resulting in spatial discontinuities when the scenes are mosaicked together (e.g. the growing season for crops near the San Francisco Bay Area peaks around April while the vegetation in adjacent scenes covering portions of the Sierra Nevada peaks later in the summer, in July or August). To rectify this issue, we harmonize the specific scenes to the same acquisition time (July-August) by replacing the April scenes with other cloud-free Landsat 5 scenes for July or August of the same year. If no Landsat 5 scenes are available during the GLS acquisition time, we choose scenes of the previous year or the year after (for 4 scenes in California).

We generate the spectral surface reflectances from the Landsat GLS data using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) atmospheric correction framework. Through the LEDAPS system, the raw data are calibrated to at-sensor radiance, converted to TOA reflectance, and then atmospherically corrected using the MODIS/6S methodology (Masek et al., 2006). Landsat surface reflectances that are derived from the TOA reflectances by the atmospheric correction routine assume that the target is lambertian and infinite and that the gaseous absorption and particle scattering in the atmosphere can be decoupled. The ancillary data used in the processing chain include gridded TOMS (Total Ozone Mapping Spectrometer) data, column water vapor from the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Protection (NCEP) reanalysis data, digital topography and NCEP surface pressure data. The dark, dense vegetation (DDV) type method is used to extract the atmospheric optical depth (AOD) directly from the imagery

(Kaufman et al., 1997). Based on the physical correlation between chlorophyll absorption and liquid water absorption, this method hypothesizes a linear relation between the shortwave-infrared (2.2 μm) surface reflectance (nearly unaffected by the atmosphere) and surface reflectance in the visible bands (Masek et al., 2006). The specific relations are derived from an analysis of data from Aerosol Robotic Network (AERONET) sites where AOD is measured directly (WWW2). The output from the LEDAPS provides surface spectral reflectances for all the available spectral bands with corresponding quality assurance (QA) flags for clouds, missing data, and an aerosol optical thickness map for the blue band. The Red and Near-infrared (NIR) spectral bands are used in our study. We also compute the normalized difference vegetation index (NDVI) from the Red and NIR bands.



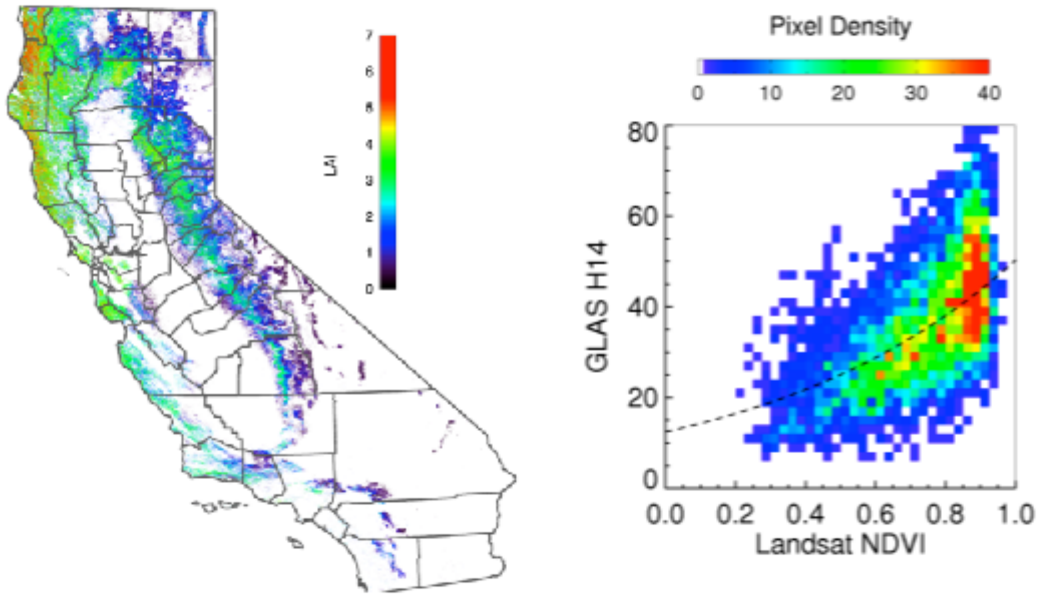
(a)

(b)

Figure 3.4: (a) Forest land classes from the NLCD 2006. (b) Sub-ecoregions map obtained from the USDA forest department. A total of 147 sub-ecoregions with forests are utilized in this analysis (http://www.fs.fed.us/r5/projects/ecoregions/preface_main.htm). The sub-ecoregions are based on an eight class ecoregions map obtained from the same source.

3.5.1.1.2 Landsat Derived LAI

We used the Landsat Red and NIR bands (Section 3.5.1.1.2) to retrieve Leaf Area Index (LAI) for all the 45 GLS scenes (Figure 3.5). A physically-based algorithm, based on the theory of canopy spectral invariants was implemented to retrieve LAI (Ganguly et al., 2012). The algorithm provides a computationally efficient way of parameterizing the Bidirectional Reflectance Factor (BRF) as a function of spatial resolution and wavelength. A reflectance model inversion is performed and operated through minimizing a merit function, which yields a value for LAI by minimizing the summed differences between simulated and measured reflectances for the Red and NIR bands.



(a)

(b)

Figure 3.5: (a) A 30-m Leaf Area Index map derived from the Landsat Global Land Survey (GLS) 2005 data. The formulation of the algorithm is described in Ganguly et al., 2012. The NLCD 2006 map is also utilized to drive the LAI retrieval process for different land cover types. (b) Density scatter plot between GLAS maximum canopy height (H14) and Landsat NDVI nearest to the GLAS center locations. The total number of sample points is 8,196. The fitted model is "H14=exp{2.523+1.395*log(NDVI)}" and the RMSE is 32.011 with a correlation of 0.48.

3.5.1.1.3 ICESat GLAS Data

The ICESat GLAS echo waveform data is referenced as the GLAS/ICESat L1A Global Altimetry Data (or GLA01). Associated with the GLA01 product, there exists a GLAS/ICEsat L2 Global Land Surface Altimetry Data (GLA14) product that provides the geo-location of each footprint along with other parameters derived from the waveform, such as the signal beginning and the echo energy peaks. In this study all the available GLA01 and GLA14 data for the time period from 2003 to 2007 are used to estimate the canopy height. Data after November 2007 are not used because of a decline in the Laser power of the GLAS instrument (Lefsky et al., 2010).

3.5.1.2 Land cover map

We use a high resolution land cover map and a sub-ecoregion map to identify the forest pixels and classify them by homogeneous eco-climatic zones.

3.5.1.2.1 NLCD Land cover

The National Land Cover Database (NLCD) 2006 provides accurate and consistent land cover at a spatial resolution of 30 m for the conterminous United States and is an updated version of NLCD 2001 (WWW3). Three land cover categories in the NLCD map, deciduous forest, evergreen forest, and mixed forest, are used in our analysis to identify forest pixels in California (Figure 3.4a).

3.5.1.2.2 Sub-ecoregion Map

A map of the ecological subregions of the conterminous US is available at 1:500,000 to 1:1,000,000 scale from the US Forest Service (USFS) ECOMAP Team (WWW5). The attributes in the map include province, section and subsection of the ecoregions, and are consistent with the terminologies used for the USFS inventory records (Figure 3.4b). A total of 147 sub-ecoregions fall in California and are used in this analysis.

3.5.1.3 Forest Inventory Data

The Forest Inventory and Analysis (FIA) Program of the U.S. Forest Service measures and records information about forests at well-designed ground plots (a standard FIA base plot includes 1 sample location per 6,000 acres roughly). In this study, we gathered tree measurements of aboveground biomass (DRYBIO), height (ACTUALHT) and the number of trees per acre (TPA) from the Forest Inventory and Analysis database (FIADB - Version 4.0) for each plot within California (WWW6) for the time period 2001-2007.

In addition to providing the plot samples, the FIA also reports the status and trend of forest tree measurements at an aggregated scale categorized by state, county and sub-ecoregion and the data is compiled FIA data for the time period 2001-2009. We obtain the total aboveground biomass values at the county and sub-ecoregion level by using the FIA Analysis EVALIDator tool (version 4.01.01; WWW7).

3.5.1.4 Ancillary Aboveground Biomass Data

3.5.1.4.1 NBCD AGB Data

The Woods Hole Research Center has produced a high-resolution “National Biomass and Carbon Dataset for the year 2000” (referred to as NBCD) at a spatial resolution of 30 m. The NBCD AGB estimation is based on an empirical modeling approach that combines USDA Forest Service FIA data with high-resolution InSAR data acquired from the 2000 Shuttle Radar Topography Mission (SRTM) mission and optical remote sensing data acquired from the Landsat ETM+ sensor (WWW8). Products from both the NLCD 2001 (land cover and canopy density) and LANDFIRE (existing vegetation type) projects as well as topographic information from the USGS National Elevation Dataset (NED) are used within the NBCD AGB estimation as spatial predictor layers for canopy height and biomass. NBCD provides two biomass maps based on different allometric equations. We used the map utilizing the Jenkins et al. (2003) allometric equation for our comparative analysis (Section 3.3).

3.5.1.4.2 USFS AGB Data

Blackard et al. (2008) provided a spatially explicit dataset of AGB (referred henceforth as USFS) by interpolating the FIA plot data mapping models through geospatial predictors. The predictors included the Moderate Resolution Imaging Spectrometer (MODIS)-derived image band composites and percent tree cover, land cover proportions from the NLCD, topographic variables, monthly and annual climate parameters, and other ancillary variables (WWW9).

3.5.2 Validity of GLAS Maximum Canopy Height

3.5.2.1 Maximum Canopy Height

The signal start of the echo waveform detects the elevation of the uppermost surface of the canopy and the last signal peak is ideally attributed to the ground surface reflection (Harding & Carabajal, 2005). The maximum canopy height is calculated as the distance between the signal start and the centroid of the ground peak recorded in GLA14 waveform data (we term the maximum canopy height as H14 from now onwards). This approach is only suitable for relatively flat surfaces where the echo waveform has a pronounced return from the ground surface compared to regions with rough topographical surfaces and high slopes (Sun et al., 2008). In order to establish the validity of the maximum canopy height retrievals, the GLA14 plots are filtered according to the following steps:

(a) The seasonal time period from May to October for each year is used as a first level screening of GLAS shots to best approximate the growing season.

(b) To avoid factors pertaining to atmospheric forward scattering and signal saturation, only the cloud-free (flag corresponding to FRir_qaFlag=15 in the GLA14 product) and saturation-free (i.e. saturation index satNdx=0 in the GLA14 product) shots were considered.

(c) A data filter was applied to make sure that the difference in elevation of the ground peak derived from the waveform and the ground elevation derived from the 30-m National Elevation Dataset (NED) data is less than 100m.

(d) Additional filtering is performed by using the terrain slope gradient information calculated from the NED 30-m elevation data. A 3x3 slope gradient kernel centered over each 30-m pixel is calculated. The center location of the GLAS shot with the nearest slope gradient pixel is matched to perform the filtering. If the slope gradient value is 0.1 and higher, the nearest GLAS shot is excluded from the analysis. The GLAS shots will be analyzed in the main algorithm.

(e) The 30-m NLCD 2006 map (Section 3.5.1.2.1) is utilized to delineate forest versus non-forest pixels. The center location of the GLAS shot with the nearest NLCD pixel center coordinate is matched to classify the GLAS shot as either non-forest or forest.

(f) Using the Landsat spectral band data (Section 3.5.1.1.2), a threshold RED spectral reflectance value of 0.3 is further used for the forest-classified pixels. If a forest pixel has a RED reflectance value of 0.3 and greater, the pixel is attributed to a bare soil or a non-vegetated pixel and is discarded from our analysis.

The total number of GLAS shots for California amounts to 560,321, amongst which 40,793 are shots over pixels classified as forest by the NLCD 2006 map. Further filtering, as described in the steps above, brings the number of usable GLAS shots used in our analysis to 8,196 (Figure 3.1).

3.5.2.2 Lorey's Height

In addition to using the maximum canopy height we also used Lorey's height, which represents the basal area weighted height of all trees. Basal area weighting of tree heights increases the importance of the largest trees in a stand and in general represents the height of a stand's tallest tree (Lefsky, 2010). We compared Lorey's height with the H14 data (Section 3.5.2.1) for a total of 1,466 samples. The H14 heights are usually higher than Lorey's height, and the coefficient of determination (R2) is 0.758 between the two height data (Figure 3.6). Lorey's height over California was obtained from the GLAS dataset by the NASA JPL Center in collaboration with Dr. Lefsky from University of Colorado at Boulder.

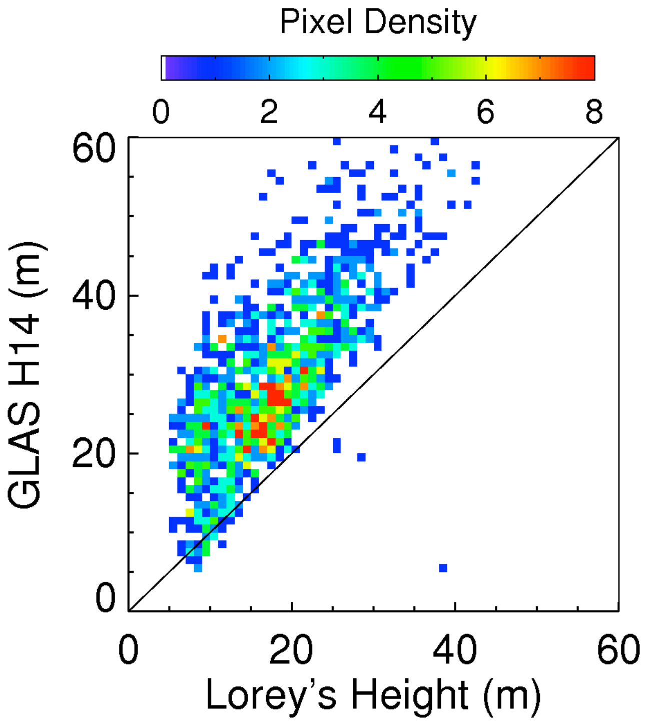


Figure 3.6: Comparison between GLAS derived maximum height (GLAS H14) and Lorey's height obtained from NASA JPL. The number of samples for GLAS H14 is obtained from Figure 3.1(a) and the Lorey's height plots are co-geolocated with the GLAS H14 data, resulting in a total of 1,466 samples. The coefficient of determination (R2) is 0.758.

3.5.3 Theoretical Relation between LAI and Maximum Tree Height

Field measurements and dynamic models show that different physiological traits of a forest are correlated with tree size following a power law. Based on a simple allometric model (West et al., 1999), the total leaf area is proportional to maximum tree height (h) such that:

$$a_l * (n^N) \propto h^3 \tag{S1}$$

where a_l is the leaf area for each leaf, n^N is the "n" number of leaves in "N" number of branches, and the product of them represents the total leaf area . Niklas and Spatz (2006) provide another relationship between canopy radius and maximum tree height such that:

$$r_{can} \propto h \tag{S2}$$

where r_{can} is the canopy radius.

Since LAI is the total leaf area divided by projected canopy area, using the proportionality rules from Eqns. S1 and S2, we can relate LAI with maximum height (Kempes et al., 2011) as:

$$LAI = \frac{a_l * n^N}{\pi * r_{can}{}^2} \propto \frac{h^3}{\pi * h^2} \propto h \tag{S3}$$

The power law between LAI and maximum tree height is a first order approximation that supports the assumption that we can obtain a near-linear relationship between Landsat derived LAI and GLAS maximum canopy height. Differences in tree density across landscapes may blur this power law - for example, in evergreen broadleaf forests, significant canopy closure will result in measured signal saturation which will affect the power law approximation. In general, the order of the power law between LAI and tree height shows variability around unity (e.g. Enquist et al., 2009 shows that LAI is proportional to $h^{0.72}$ in tropical regions). In this study, the conifer-dominated forests of California are best modeled by a simple linear relationship between tree height and LAI. We also tested nonlinear relationships between height and LAI by building an error propagation model to demonstrate the effect of using different power coefficients for height in regulating final AGB estimates (please see Section 3.5.7).

3.5.4 Development of the Height-Biomass Model

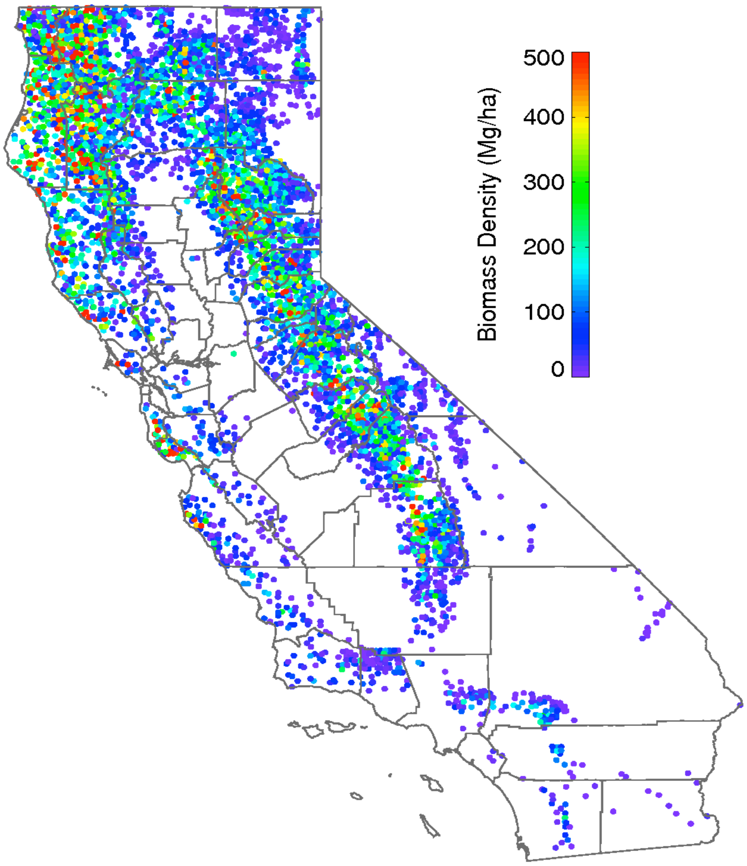


Figure 3.7: Plot level aggregated FIA aboveground biomass density estimates. The FIA plots level biomass estimates are generated from the FIADB raw data (http://apps.fs.fed.us/fiadb-downloads/datamart.html). The aboveground biomass density for a plot is a summation of individual tree biomass estimates weighted with an intrinsic factor, such that $BD_{plot} = \sum_{i=1}^{n} CARBONAG_i \times TPA_i \times 2.0$, where n is the number of trees in each plot, $CARBONAG_i$ is Carbon in the aboveground portion of a tree, TPA_i is the scaling factor resembling trees per acre, and 2.0 is a scalar factor to convert carbon to biomass in FIADB.

To develop the height biomass model we gathered data from 2,205 FIA plots in California for the period 2000 to 2007 (Section 3.5.1.3). The maximum and mean tree heights are calculated from all the trees available in each plot. The aboveground biomass density for a plot (BD_{plot}) is the summation of individual tree biomass estimates weighted by an intrinsic factor, such that

$$BD_{plot} = \sum_{i=1}^{n} CARBONAG_i \times TPA_i \times 2.0 \qquad \text{(S4)}$$

where n is the number of trees in each plot, $CARBONAG_i$ is Carbon in the aboveground portion of a tree, TPA_i is the species-specific scaling factor representing trees per acre and 2.0 is a scalar factor to convert carbon to biomass in FIADB. The spatial representation of the plot level biomass density is shown in Figure 3.7.

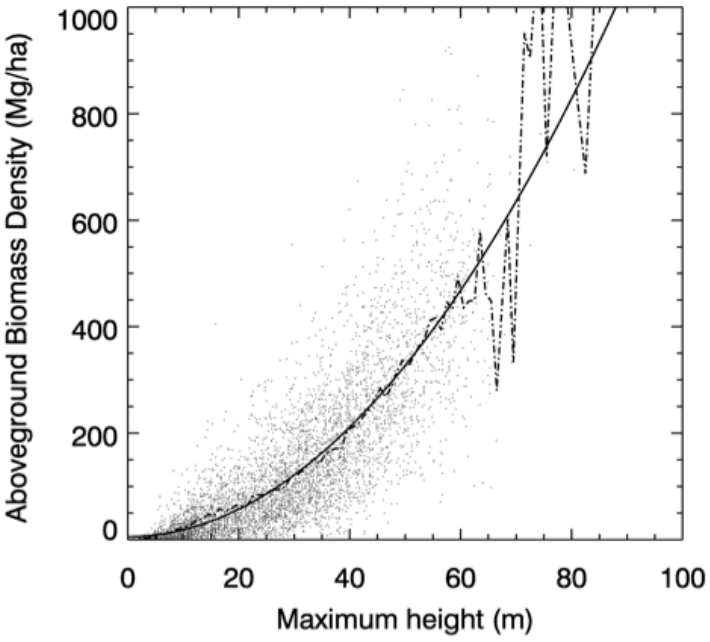


Figure 3.8: Relationship between height and aboveground estimates for FIA plots for the period 2004-2006. The plot level biomass in (Mg/ha) is calculated as in Figure 3.7. The total number of samples is 2205. The model for maximum height is $AGB = 2.387+0.135* (Hmax^2)$ and $R^2=0.642$. The dotted line is the bin average and the solid line is the fitted model.

We built the relationship between maximum canopy height squared for each plot and the total biomass using the equation in Lefsky et al. (2005). When using maximum heights the resulting model is AGB = 2.387+0.135* ($Hmax^2$) and gives a R^2 of 0.642 (Figure 3.8).

3.5.5 Variation of Biomass Density with Spatial Scales

To analyze the variation of biomass density with spatial resolution, we perform an upscaling experiment by spatially aggregating our 30-m biomass density map (ARC) for a region spanning the hardwood forests of Northern Sierra (covering an area of ~5500 square miles). We start with a majority resampling of the 30-m NLCD land cover map to resolutions that are a factor of n (where n = 1 to 20). If a pixel grid at 60-m is identified as forest, we calculate the biomass density at this grid as the mean of the 4 corresponding 30-m pixels. If the pixel is non-forest, the biomass value is set to zero, which is commonly algorithm for forest aboveground biomass estimation. In this process, we calculate the mean biomass density, the dispersion in biomass density and the forest coverage for the whole region at different resolutions (Figure 3.9). The figure shows that the forest cover remains nearly constant during the upscaling process (~0.2% relative change). The mean biomass density and the standard deviation of biomass density from the ARC map decrease monotonically at coarser resolutions, declining by ~ 10.6% at the USFS equivalent spatial resolution, since the biomass in sparse forests are dropped off during the majority resampling of land cover. This upscaling exercise confirms the hypothesis that aggregating biomass density across different spatial resolutions is theoretically incorrect and that maps of AGB densities with different spatial resolution cannot be directly compared.

3.5.6 Calculation of Regional Total Biomass

Our inter-comparison exercise between satellite-derived AGB maps and FIA is restricted to two aggregated scales – a) county level and b) sub-ecoregion level. For each of county sub-ecoregion the AGB density values are converted to “total biomass” in Mega tons (M ton). These aggregated scales are also used to guide the FIA data plot design strategy. FIA total biomass data for the county and sub-ecoregions are readily obtained from the FIA EVALIDator tool (Table 3.4 and Section 3.5.1.3).

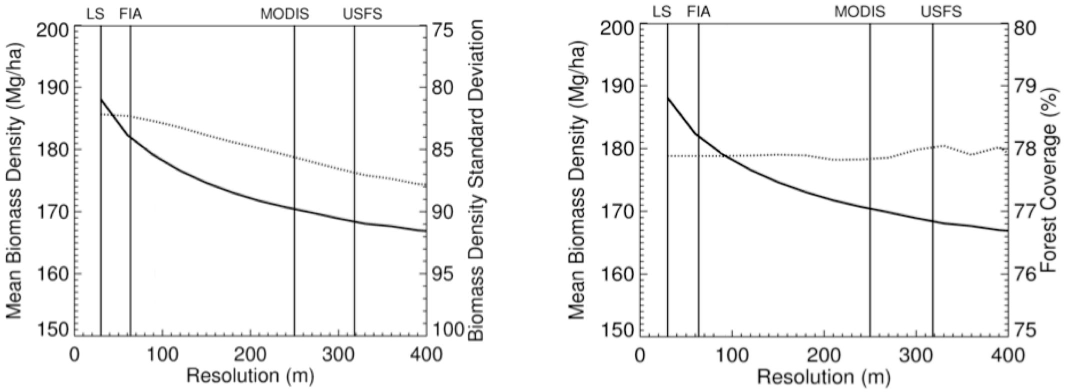


Figure 3.9: Variation of ARC biomass density mean and standard deviation with changes in spatial resolution. The NLCD map was resampled to the resolutions noted along the x-axis. For each resolution, the biomass numbers are calculated for the region of interest. The region of interest spans a wide region of hardwood forests in California covering an area of ~5500 square miles. If a pixel represented by the scale of interest is classified as forest, the biomass is calculated as the average from the total number of native 30m pixels; for a pixel that is classified as non-forest the biomass is set to zero.

Table 3.4: Datasets used in this research

Data	**Source**	**Time Period**	**Notes**
Remote Sensing Data			
Landsat GLS 2005	USGS/NASA	2004-2006	Surface Reflectances were calculated using the LEDAPS processing framework
Landsat LAI	NASA ARC	2004-2006	
Geoscience Laser Altimeter System (GLAS)	ICESat	2003-2007	The waveform data from GLA14 product was utilized
Lorey’s Height	JPL/ Lefsky et al., 2010	2003-2007	This data was provided by JPL

The county/sub-ecoregion total biomass from the satellite-derived maps is calculated as:

$$AGB_{Total} = \left(\sum_{n=1}^{P_n} BD_n \right) \times Area \qquad \text{(S5)}$$

where Area is the pixel area, *BD* is the Biomass Density, P_n is the number of forest pixels within the respective sub-ecoregions based on USFS sub-ecoregion map. The county total biomass is calculated in the same manner, in which case P_n is the number of forest pixels within the respective county. Figures 3.3b,c and Figure 3.10 show the total aboveground biomass value for counties and sub-ecoregions within California. The variance of total AGB among county and sub-ecoregions is significant – with AGB ranging from less than 0.1 M ton to greater than 200 M ton (Figure 3.3b & c).

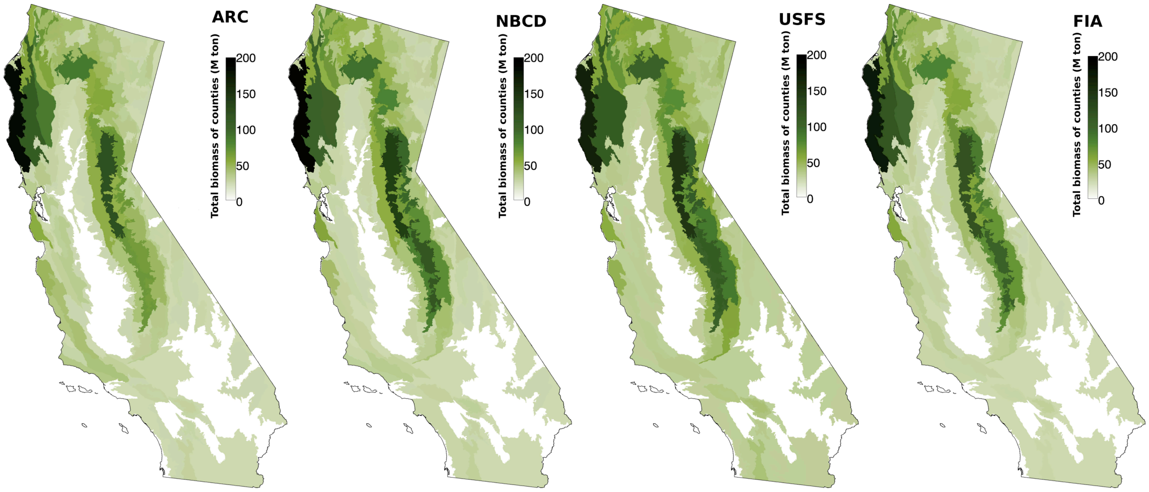


Figure 3.10: Total aboveground biomass (in M ton) estimates at sub-ecoregion level from the ARC, NBCD, USFS and FIA biomass density maps. The sub-ecoregion map is showed in Figure 3.4(b). Estimation of total biomass is explained in Section 3.5.6. The sub-ecoregion total biomass from the satellite-derived maps is calculated as $AGB_{Total} = \left(\sum_{n=1}^{P_n} BD_n\right) \times Area$, where Area is the pixel area, BD is the Biomass Density, and P_n is the number of pixels for the respective sub-ecoregion.

Figure 3.11 compares the differences between the various remote sensing derived AGB and the FIA estimates for each sub-ecoregion. We analyze the differences in spatial patterns of aggregated AGB values, by stratifying the sub-ecoregions by biomass class intervals and calculating statistical metrics as presented in Table 3.5 and Table 3.6. Most forests in California are located in five sub-ecoregion (Table 3.5) representing 99.1% of the total forest AGB. The ARC total AGB values compare best to the FIA estimates for the dominant forested ecosystem provinces (M261 and M263 encompass approximately 50% of the total area of the sub-ecoregions) while they show a larger difference w.r.t. FIA values for the rest of the ecoregions. To supplement the analysis further, Table 3.6 provides a coefficient of determination matrix for sub-ecoregions grouped by total biomass ranges, illustrating the disparities that can be seen when we go from low to high biomass. The correlation between different AGB maps and FIA improves when we pool all the sub-ecoregions together and for the sub-ecoregions with higher biomass (>10 M ton) but goes down significantly when we pool sub-ecoregions that cover the intermediate (1 to 10 M ton) and lower biomass ranges (<1 M ton). The ARC and JPL AGB values still show higher correlation across the whole biomass range. The similar correlation can be explained based on the fact that the two AGB estimation approaches depend on similar satellite-derived input metrics.

The calculation of the total biomass aids the analysis as described in Section 3.2 and 3.3.

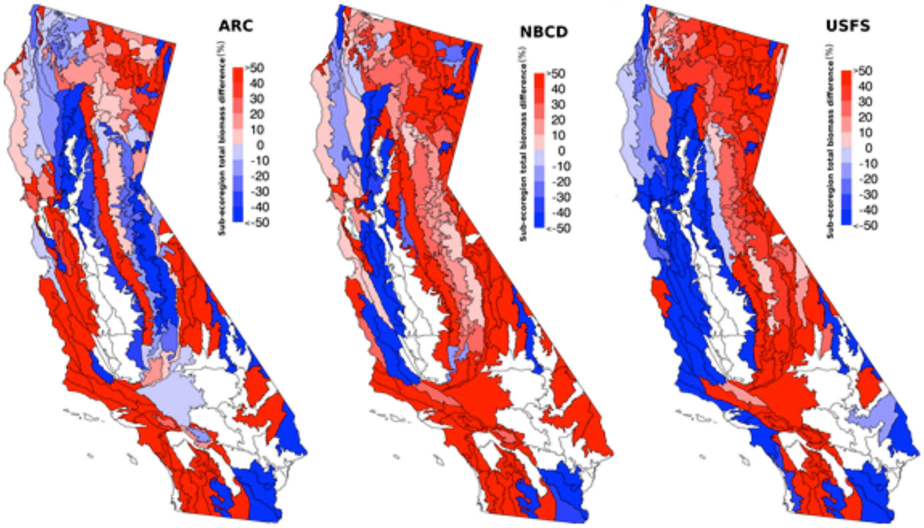


Figure 3.11: Difference between the satellite and the FIA-derived total biomass estimates by sub-ecoregion. The difference is estimated as $\Delta=((B_{RS}-B_{FIA})/B_{FIA})x100$, where B is the sub-ecoregion total aboveground biomass.

Table 3.5: Total biomass estimates (in M ton) and relative difference between total biomass and FIA estimates (in parenthesis) by ecoregion for California. The relative difference is calculated as ((RS – FIA)/FIAx100), where RS refers to ARC/NBCD/USFS estimates.

Region **(No of sub-ecoregions)**	**ARC**	**NBCD**	**USFS**	**FIA**
M261 (67)	1270.3 (-8.0%)	1652 (19.6%)	1760 (27.5%)	1381
263 (7)	282.7 (0.3%)	273.5 (-2.9%)	235.4 (-16.4%)	281.7
261 (14)	110.3 (57.8%)	93.5 (33.8%)	53.9 (-23.0%)	69.9
M262 (22)	98.3 (82.9%)	63 (16.8%)	66 (24.3%)	54
262 (6)	1.2 (-76.6%)	6.3 (23.1%)	0.1 (-98.5%)	5.1
Others (31)	24.4 (50.0%)	59.62 (221.4%)	37.8 (145.1%)	15.4

Table 3.6: Coefficient of determination matrix for sub-ecoregions for different total biomass ranges as calculated from Figure 3.10.

(a) All 147 sub-ecoregions

Model	NBCD	USFS	FIA
ARC	0.876	0.928	0.940
NBCD		0.970	0.966
USFS			0.938

(b) Biomass over 10 M ton (N = 41)

Model	NBCD	USFS	FIA
ARC	0.893	0.818	0.913
NBCD		0.950	0.946
USFS			0.901

(c) Biomass from 1 to 10 M ton (N = 54)

Model	NBCD	USFS	FIA
ARC	0.380	0.104	0.418
NBCD		0.650	0.361
USFS			0.251

(d) Biomass less than 1 M ton (N = 52)

Model	NBCD	USFS	FIA
ARC	0.046	0.010	0.046
NBCD		0.166	0.032
USFS			0.010

3.5.7 Uncertainty Modeling

In order to assess the uncertainty of our AGB density estimates, we implement a simple error propagation model assuming that the errors in the distribution of forest biomass can be random or systematic in nature and can include the following: (a) observation errors associated with the uncertainty in estimates of maximum height derived from GLAS Lidar, spectral surface reflectance and resulting LAI from Landsat and NLCD land cover classification; (b) prediction errors associated with LAI to maximum height prediction and from height to biomass. In building the uncertainty model, we assume that the NLCD 2006 classification accuracy for forest classes is acceptable and hence we do not introduce land cover uncertainties. Additionally, there are no spatial sampling errors as the native resolution of the land cover map is the same as Landsat. We realize that the pixel specific accuracy in maximum height values directly depends on the LAI-to-Height prediction model, which implicitly also accounts for errors in GLAS derived height metrics (c.f. Section 3.5.2). The pixel specific error propagation model can be written in a functional form where the variation in AGB is related to the variation in LAI such that

$$\Delta B(p)^2 = \left(\frac{\partial B}{\partial LAI}\right)^2 \times \Delta LAI(p)^2 \tag{S6}$$

where $\Delta B(p)$ is the uncertainty of AGB density for the pixel p, B is the pixel AGB density and $\Delta LAI(p)$ is the uncertainty in LAI for the pixel p.

Now, from Section 3.5.2 and Figure 3.1, the Height-LAI functional relationship can take the form:

$$H(p) = e_1 + e_2 \cdot LAI(p)^{\alpha} \qquad \because \alpha = [0,2] \tag{S7}$$

where $H(p)$ is the predicted height for pixel p and α regulates the polynomial degree of the equation. e_1 and e_2 are the linear constants.

Also, from Section 3.5.4, we can write the Height to Biomass relationship as

$$B(p) = a + b \cdot H(p)^2 \tag{S8}$$

We can express Eqn. (S6) as

$$\Delta B(p) = \left(\frac{\partial B}{\partial H}\right) \times \left(\frac{\partial H}{\partial LAI}\right) \times \Delta LAI(p) \tag{S9}$$

From Eqn. (S7), we get $\frac{\partial H}{\partial LAI} = \alpha \cdot e_2 \cdot LAI(p)^{\alpha-1}$.

and from Eqn. (S8), we get $\frac{\partial B}{\partial H} = 2 \cdot b \cdot H(p) = 2 \cdot b \cdot \left(e_1 + e_2 \cdot LAI(p)^{\alpha}\right)$.

Substituting $\frac{\partial H}{\partial LAI}$ and $\frac{\partial B}{\partial H}$ in Eqn. (S9), we have

$$\Delta B(p) = 2 \cdot b \cdot \left(e_1 + e_2 \cdot LAI(p)^{\alpha}\right) \times \left(\alpha \cdot e_2 \cdot LAI(p)^{\alpha-1}\right) \times \Delta LAI(p) \tag{S10}$$

The expression in Eqn. (S10) now explicitly incorporates the input errors in pixel retrieved LAI and predicted height values as a function of LAI.

In order to map the uncertainties in AGB densities, we introduce two scenarios:

Case (A): $\Delta LAI(p)$ = 0.5 (based on the accuracy of the standard MODIS LAI product is 0.5 under stage 2 validation efforts) i.e. there is a uniform bias of 0.5 units of LAI for all forest pixels. If the LAI varies from 1 to 7.1, the percent uncertainty, calculated as $\frac{\Delta LAI(p)}{LAI(p)} \times 100$, will vary from ~7% (LAI=7.1) to 50% (LAI=1).

Substituting the value of $\Delta LAI(p)$ in Eqn. (S10), we simulate $\Delta B(p)$ for four different values of α (α=0.5, 1, 1.5 and 2) and calculate the pixel specific percent uncertainty in AGB as $\mu_B = \frac{\Delta B(p)}{B(p)} \times 100$. The variation in α accounts for the bias introduced for model selection of estimating height from LAI.

Case (B): Same as *Case (A)* but with $\Delta LAI(p)$ = 1 i.e. there is a uniform bias of 1 unit of LAI for all forest pixels. In this case the percent uncertainty in LAI varies from 14.08% (LAI=7.1) to 100% (LAI=1).

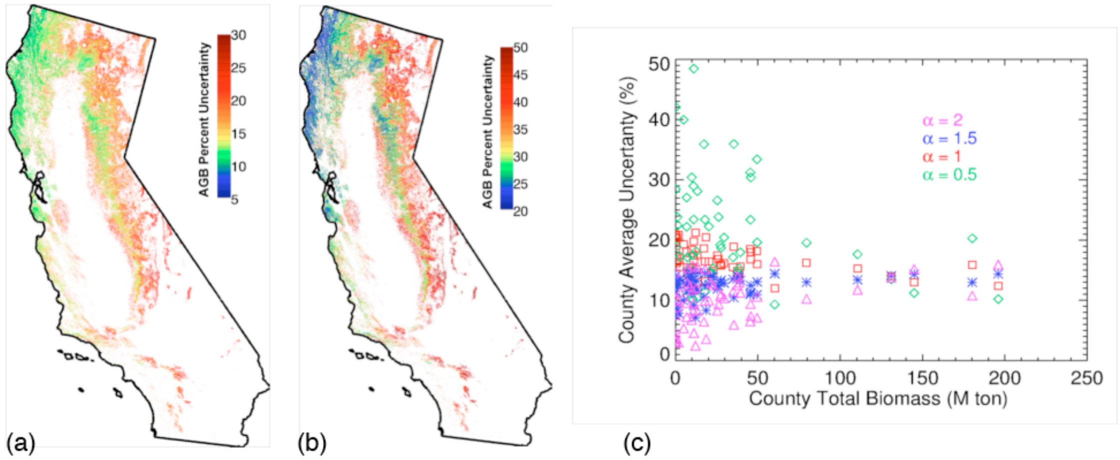


Figure 3.12: (a) Percent uncertainty in AGB density estimates. The uncertainty value is calculated by propagating a uniform bias in LAI =0.5 and by choosing a linear fitting model (α=1) for estimating height from LAI (refer to 3.5.7 for more details). (b) Same as (a) but for a bias in LAI=1. (c) County average uncertainty as a function of county total biomass for different values of α.

Figure 3.12 (a) and (b) show example scenarios for Case (A) and Case (B) and for α equal to 1. In order to demonstrate the effect of four different polynomial kernels, we compute the county level mean biomass uncertainty (averaging the pixel level AGB density uncertainty values within the specific county) for different values of α. Figure 3.12 (c) shows the county average percent uncertainty with respect to county total biomass – it is noted that for α=1, 1.5 and 2, the county estimates were found to be constrained within 0-20% of the total biomass obtained from averaging the pixel AGB density values except for α=0.5, where the percent uncertainty ranges from 0-50% for low values of total biomass. In a separate exercise, we also calculate the uncertainty numbers at a sub-ecoregion level by grouping the sub-ecoregions by total biomass classes (Table 3.7) – we portray three scenarios with varying uncertainties in LAI (uniform bias of 0.5, 1 and 1.5) and for different power coefficients that control the LAI-height equation. The percent uncertainties are constrained between 11 to 38% for both *Case 1* and *Case 2* and for α=1 and 1.5 (see Table 3.7), however, the uncertainties are significantly higher in lower biomass regions when the input bias in LAI is 1.5 (for example if the LAI of a given pixel is 0.5 and the observed bias is 1.5 LAI units, the percent uncertainty in LAI is ~300%).

This exercise shows that the uncertainty in the AGB estimates can be constrained within an acceptable range but will vary depending on the scale at which it is reported.

Table 3.7: Uncertainty in sub-ecoregion total biomass estimates for different case scenarios as described in Section 3.5.7. The total biomass ranges and the sub-ecoregions that fall under these ranges are obtained from Figure 3.3(c).

Total Biomass Range (M Ton)	**Uncertainty in Sub-ecoregion Biomass (%)**		
Case 1			
$\Delta LAI(p)$=0.5	α=0.5	α=1.0	α=1.5
>20	14.83	16.90	13.49
4.5 - 20	16.95	25.40	12.11
<4.5	17.77	33.05	11.12
Case 2			
$\Delta LAI(p)$=1.0	α=0.5	α=1.0	α=1.5
>20	36.35	31.67	28.86
4.5 - 20	55.59	36.46	25.88
<4.5	73.31	38.34	23.89
Case 3			
$\Delta LAI(p)$=1.5	α=0.5	α=1.0	α=1.5
>20	57.90	50.51	46.32
4.5 - 20	89.56	58.53	41.54
<4.5	119.11	617.24	383.27

References

Abshire, J. B., et al. (2005). Geoscience Laser Altimeter System (GLAS) on the ICESat Mission: On-orbit measurement performance. *Geophysical Research Letters*, *32*(21), doi:10.1029/2005GL024028.

Asner, G. P., et al. (2010), High-resolution forest carbon stocks and emissions in the Amazon, *Proceedings of the National Academy of Sciences*, *107*(38), 16738 -16742, doi:10.1073/pnas.1004875107.

Blackard, J. A., Finco, M. V., Helmer, E. H., Holden, G. R., Hoppus, M. L., Jacobs, et al. (2008). Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information. *Remote Sensing of Environment*, *112*(4), 1658-1677. doi:16/j.rse.2007.08.021.

Baccini, A., N. Laporte, S. J. Goetz, M. Sun, and H. Dong (2008), A first map of tropical Africa's above-ground biomass derived from satellite imagery, *Environmental Research Letters*, *3*, 045011, doi:10.1088/1748-9326/3/4/045011.

Baccini, A., S. J. Goetz, N. Laporte, M. Sun, and H. Dong (2011), Reply to Comment on "A first map of tropical Africa"s above-ground biomass derived from satellite imagery', *Environmental Research Letters*, *6*, 049002, doi:10.1088/1748-9326/6/4/049002.

Baccini, A., et al. (2012), Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps, *Nature Climate Change*, doi: 10.1038/NCLIMATE1354.

Boudreau, J., et al. (2008), Regional aboveground forest biomass using airborne and spaceborne LiDAR in Québec, *Remote Sensing of Environment*, *112*, 3876-3890, doi:10.1016/j.rse.2008.06.003.

Canadell, J. G., et al. (2007), Contributions to accelerating atmospheric CO2 growth from economic activity, carbon intensity, and efficiency of natural sinks, *Proceedings of the National Academy of Sciences*, *104*(47), 18866 -18870, doi:10.1073/pnas.0702737104.

Drake, J. B., R. O. Dubayah, R. G. Knox, D. B. Clark, and J. Blair (2002), Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest, *Remote Sensing of Environment*, *81*, 378-392, doi:10.1016/S0034-4257(02)00013-5.

Enquist, B. J., West, G. B., and Brown, J. H. (2009), Extensions and evaluations of a general quantitative theory of forest structure and dynamics, *Proceedings of the National Academy of Sciences*, *106*, 7046-7051.

Ganguly, S., et al. (2012), Generating global leaf area index from Landsat: Algorithm formulation and demonstration, *Remote Sensing of Environment*, http://dx.doi.org/10.1016/j.rse.2011.10.032.

Ganguly, S., M. Schull, A. Samanta, N. Shabanov, C. Milesi, R. Nemani, Y. Knyazikhin, and R. Myneni (2008), Generating vegetation leaf area index earth system data record from multiple sensors. Part 1: Theory, *Remote Sensing of Environment*, *112*, 4333-4343, doi:10.1016/j.rse.2008.07.014.

Gibbs, H. K., S. Brown, J. O. Niles, and J. A. Foley (2007), Monitoring and estimating tropical forest carbon stocks: making REDD a reality, *Environmental Research Letters*, *2*, 045023, doi:10.1088/1748-9326/2/4/045023.

Goetz, S. J., et al. (2009), Mapping and monitoring carbon stocks with satellite observations: a comparison of methods, *Carbon Balance Manag.*, *4*, 2, doi:10.1186/1750-0680-4-2.

Gurney, K. R., and L. Raymond (2008), Targeting deforestation rates in climate change policy: a "Preservation Pathway" approach, *Carbon Balance Manag.*, *3*, 2, doi:10.1186/1750-0680-3-2.

Gutman, G., Byrnes, R., J. Masek, S. Covington (2008) Towards Monitoring Changes at a Global The Global Land Su, *Photogrammetric Engineering & Remote Sensing 7*.

Harding, D. J., and C. C. Carabajal (2005), ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure, *Geophys. Res. Lett.*, *32*, 4 PP., doi:200510.1029/2005GL023471.

Houghton, R. A. (2005), Aboveground Forest Biomass and the Global Carbon Balance, *Global Change Biology*, *11*(6), 945-958, doi:10.1111/j.1365-2486.2005.00955.x.

Houghton, R. A., et al. (2000), Annual fluxes of carbon from deforestation and regrowth in the Brazilian Amazon, *Nature*, *403*(6767), 301-304, doi:10.1038/35002062.

IPCC, C. C. (2007), The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*, *996*, 2007.

Kaufman, Y.J., et al. (1997). The MODIS 2.1-μm channel-correlation with visible reflectance for use in remote sensing of aerosol, *Geoscience and Remote Sensing, IEEE Transactions on 35,* 1286–1298.

Keith, H., B. Mackey, S. Berry, D. Lindenmayer, and P. Gibbons (2010), Estimating carbon carrying capacity in natural forest ecosystems across heterogeneous landscapes: addressing sources of error, *Global Change Biology*, *16*(11), 2971-2989, doi:10.1111/j.1365-2486.2009.02146.x.

Lefsky, M. A. (2010), A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophysical Research Letters*, *37*, 15401.

Lefsky, M. A., D. Harding, W. Cohen, G. Parker, and H. Shugart (1999), Surface Lidar Remote Sensing of Basal Area and Biomass in Deciduous Forests of Eastern Maryland, USA, *Remote Sensing of Environment*, *67*, 83-98, doi:10.1016/S0034-4257(98)00071-6.

Lefsky, M. A., et al. (2005), Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters*, *32*.

Lu, D. (2006), The potential and challenge of remote sensing-based biomass estimation, *International Journal of Remote Sensing*, *27*(7), 1297-1328.

Masek, J.G., et al. (2006) A Landsat surface reflectance dataset for North America, 1990-2000, *Geoscience and Remote Sensing Letters, IEEE 3*, 68–72.

Miles, L., and V. Kapos (2008), Reducing Greenhouse Gas Emissions from Deforestation and Forest Degradation: Global Land-Use Implications, *Science*, *320*(5882), 1454 -1455, doi:10.1126/science.1155358.

Kempes, C. P., G. B. West, K. Crowell, and M. Girvan (2011), “Predicting Maximum Tree Heights and Other Traits from Allometric Scaling and Resource Limitations.” *PLoS ONE*, *6*(6), e20551. doi:10.1371/journal.pone.0020551.

Niklas, K. J., and H-C. Spatz (2006), “Allometric theory and the mechanical stability of large trees: proof and conjecture.” *American Journal of Botany*, *93*(6), 824 -828, doi:10.3732/ajb.93.6.824.

Saatchi, S. S. et al. (2011), Benchmark map of forest carbon stocks in tropical regions across three continents, *Proceedings of the National Academy of Sciences*, *108*(24), 9899 -9904, doi:10.1073/pnas.1019576108.

Saatchi, S. S., R. A. Houghton, R. C. Dos Santos Alvalá, J. V. Soares, and Y. Yu (2007), Distribution of aboveground live biomass in the Amazon basin, *Global Change Biology*, *13*(4), 816-837, doi:10.1111/j.1365-2486.2007.01323.x.

Sun, G., Ranson, K. J., Kimes, D. S., Blair, J. B., & Kovacs, K. (2008). Forest vertical structure from GLAS: An evaluation using LVIS and SRTM data. *Remote Sensing of Environment*, *112*(1), 107-117. doi:16/j.rse.2006.09.036.

West, Geoffrey B., James H. Brown, and Brian J. Enquist. 1999. A general model for the structure and allometry of plant vascular systems. *Nature*, *400*(6745), 664-667. doi:10.1038/23251.

UNFCCC (http://unfccc.int/methods_and_science/lulucf/items/4123.php)

WWW1: http://landsat.usgs.gov/science_GLS2005.php

WWW2: http://ledaps.nascom.nasa.gov/docs/pdf/SR_productdescript_dec06.pdf

WWW3: http://www.mrlc.gov/nlcd_2006.php

WWW4: http://frap.cdf.ca.gov/projects/frap_veg/index.html

WWW5: http://svinetfc4.fs.fed.us/clearinghouse/other_resources/ecosubregions.html

WWW6: http://apps.fs.fed.us/fiadb-downloads/datamart.html

WWW7: http://apps.fs.fed.us/Evalidator/tmattribute.jsp

WWW8: http://svinetfc4.fs.fed.us/clearinghouse/rastergateway/biomass/

WWW9: http://fsgeodata.fs.fed.us/rastergateway/biomass/

CHAPTER 4

FIRST GLOBAL STAND AGE MAP FOR PRIMARY FORESTS BY USING REMOTE SENSING DATA

Abstract

Disturbance plays an important role in terrestrial carbon storage and dynamics, especially for forests whose large carbon pool –aboveground biomass (AGB) - could be dramatically reduced by disturbances. The stand age of large regions represents the frequencies of disturbance and is a key factor in analysis and modeling of carbon cycle. Characterizing global forest stand age is still challenging in regions without abundant long-term inventory data. This study, for the first time, I generated a stand age map for global primary forests based on remote sensing-derived AGB and net primary production (NPP). I built a theoretical framework by coupling a theoretical carbon accumulation function with the probability density functions of disturbance occurrences, thus linking potential biomass, current biomass, net carbon fluxes into biomass, and disturbance frequency. I compared my stand age estimates against another stand age map based on interpolation of extensive inventory data in North American and found overall agreement but some important regional differences. Uncertainties in my approach are not trivial, and I explored the sensitivity of estimated stand age to variations in model parameters and input variables.

4.1 Introduction

Global forests are an important component of the terrestrial carbon cycle, and contribute to a net carbon sink of about 1 Pg carbon per year, and net carbon balance tends to be positive (i.e. indicating net uptake of carbon from the atmosphere) in established and regrowing forests but negative in cases of deforestation and land-use change (Canadell and Raupach, 2008; Pan et al., 2011). Compared to other biomes, forests store more carbon (~80% living biomass) in aboveground biomass (AGB), which is also usually the most rapidly changing carbon pool (Houghton, 2005; Fahey et al., 2009). The carbon stored in forest AGB is frequently released to atmosphere by natural and anthropogenic disturbances (Achard et al., 2002). These two types of disturbances are the primary mechanisms affecting the change from carbon sinks to sources in ecosystem (Baldocchi, 2008). It is therefore of key importance to understand the timing of forest disturbances and their effect on current and potential carbon stored in terrestrial ecosystems (Gurney and Raymond, 2008).

Ecosystem carbon storage is controlled by two types of processes: (1) internal processes, including photosynthesis, respiration, decomposition, and natural mortality, through which carbon stored in ecosystems is gradually accumulated or consumed; (2) natural or anthropogenic disturbances on ecosystem structure and function (Tilman, 1985), in which sudden mortality removes the carbon from live biomass (Huston, 2004). Both processes play crucial roles in carbon storage, but with different mechanisms and at various spatio-temporal scales. Compared to internal processes, disturbances are generally temporally stochastic with specific spatial range (O’Halloran et al. 2011) linked to the underlying human or natural forcings, such as land-use change, wild fires, hurricanes, and insect outbreaks (Bender et al., 2010; Kurz et al., 2008; Turetsky et al., 2010). Disturbances also differ in severity, frequency, magnitude, and their impacts on forest structure and function (Westerling et al., 2011). It is necessary to account for carbon accumulation and disturbance in assessing and modeling current and potential forest carbon storage, since regional carbon storage capacity -- the carbon mass stored in an ecosystem under a dynamic equilibrium (sometimes referred to as carbon storage capacity) -- is always lower than the sum of individual tree potential maximum biomass due to disturbance (Luo and Weng, 2010). In principle, it should be possible to estimate carbon storage capacity by first approximating potential maximum biomass via remotely sensed net primary production (NPP) and then reducing potential biomass as a function of the local disturbance regime (Keith et al., 2009; Keith et al., 2010).

Measurements of disturbances are a major source of uncertainty in global carbon cycle research (Mildrexler et al., 2007). Remote sensing techniques expanded measurements of disturbance events from local (Lentile et al., 2006) to global scales (Running, 2008) in the past several decades. The data from remote sensors is used to detect active disturbance (Ichoku et al., 2003) and ecosystem changes in post-disturbance spectral response (Chambers et al., 2007). Passive remote sensing is useful for determining near-past disturbances, but cannot directly capture the influence of disturbances occurring well before the remote sensing era, especially for quickly-recovering forests (Turner, 2010). It may be possible, though, to use tree height estimates from active remote sensing approaches (such as Light Detection And Ranging, LiDAR) to infer past disturbances in ecosystems in which height is related to prior disturbance intensity (Sun et al. 2008; Goetz et al. 2010). Unfortunately, LiDAR data are sparse and point-based and therefore cannot provide wall-to-wall estimates of stand age, particularly because stochastic disturbances occur at specific regions, leading to a patchwork mosaic of stand ages and tree heights on landscape scale (Weir, et al., 2000).

To avoid the uncertainty from the spatial and temporal contingency described above, it is necessary to explore the general natural disturbance regimes at landscape scales by averaging the seral stages or stand ages over a sufficiently large area at landscape equilibrium (Turner et al., 1993). Here I define stand age as the mean age of all trees inside a stand. This definition may not directly apply to ecosystems in which individual trees die, but the stand itself persists indeterminately, such as some tropical forests. The concept of landscape equilibrium is generally applied to describe a region where the spatial pattern of stand age is aggregated up from the mosaic of stand age and biomass created by past disturbances (Connell and Sousa, 1983;

DeAngelis and Waterhouse, 1987). By aggregating previous disturbances, the overall disturbance regime in landscape equilibrium can be characterized by a spatial parameter (disturbance severity) and a temporal parameter (mean disturbance interval, MDI) (Smithwick, et al. 2007; Turner et al., 1993). Disturbance severity, which is the proportion of trees and their propagules killed by one disturbance within the landscape (Oliver, et al., 1990), directly links to the decline of biomass during disturbance. The frequency of disturbance occurrence is represented by MDI, which is mean interval time between disturbances. MDI is the same as the mean stand age in a landscape, given that the severity is counted as the death rate of individual trees during one disturbance. Therefore, the pattern of stand age at the landscape scale represents the natural disturbance regime under prevailing environmental conditions (Franklin et al., 2002).

To date, forest inventories mostly measure the tree stand age from tree rings or a combination of remote sensing and stand inventories to produce wall-to-wall regional stand age maps (Chen et al., 2003; Pan et al., 2011). Characterizing the global forest stand age is still challenging in regions without remote sensing disturbance detection and abundant inventory data. In this study, for the first time, I generated a stand age map for global primary forests based on the latest remote sensing-derived AGB and NPP estimation. I incorporated a carbon accumulation equation and disturbance occurrence probability in a framework to describe the analytical linkage among the carbon storage, carbon influx and disturbance regimes. I calculated the carbon residence time (CRT) from maximum AGB and mean NPP over the landscape and then introduced a spatial-to-temporal conversion to link AGB to the frequency of disturbance on the equilibrium landscape (Van Wagner, 1978). Hence I solve for stand age from spatial patterns of AGB in an equilibrium landscape based on simultaneous equations of carbon accumulation and natural disturbance occurrence.

4.2 Methodology and Data

4.2.1 Carbon Accumulation Model

Forests gain biomass from NPP but lose carbon due to turnover, characterized here as the mean CRT in biomass. The difference between NPP and turnover determines how much biomass can be accumulated by a forest, so derivation of current biomass is expressed as,

$$\frac{d\,B(t)}{d\,t} = NPP - \frac{B(t)}{\tau} \qquad \text{Eq. 4.1}$$

where *NPP* (gC/m^2/yr) is the mean NPP for each grid; *B(t)*(gC/m^2) is the biomass at time *t* (yr), and τ (yr) is CRT (Figure 4.1a). Similar carbon cycle formulation is widely used in ecosystem modeling, such as CENTURY model (Parton et al., 1993) and TECO (Luo et al., 2003)

In the absence of disturbance, for example, in old growth forests, forest biomass eventually reaches its potential (maximum) capacity when its carbon gain through NPP is balanced by its carbon loss via natural mortality (Stephenson and Mantgem, 2005). Such a relationship can be described by an idealized equation,

$$NPP - \frac{B_p}{\tau} = 0 \qquad \text{Eq. 4.2}$$

where, B_p(gC/m^2) is potential maximum biomass under local climate condition. Clearly, the CRT (τ) can be estimated given NPP and biomass. The mean CRT is firstly calculated from maximum biomass and mean NPP within landscape and used later as the parameter for stand age estimation.

In reality, the accumulation of biomass may be interrupted by the occurrence of disturbances. Thus the potential biomass without disturbances may never be achieved at large spatial scales. The current carbon pool is the summation of the carbon increment (influx) since last disturbance and the legacy carbon left after last disturbance. If the carbon influx rate is assumed to be constant (details addressed in NPP estimation), the carbon pool could be estimated by the integration of Eq. 4.1,

$$B_t = (B_0 \times s) \times e^{\frac{-t}{\tau}} + NPP \times \tau \times (1 - e^{\frac{-t}{\tau}}) \qquad \text{Eq. 4.3}$$

where, B_t(gC/m^2) is the current carbon content, B_0(gC/m^2) is previous biomass before disturbance, thus the legacy carbon right after a disturbance event is given by B_0 timed the disturbance severity (s) , and t(yr) is the time counted since the last disturbance event (Figure 4.1b).

4.2.2 Occurrence Probability of Disturbances

The stochastic nature of disturbances poses new difficulties in describing forest biomass dynamics. For example, information about disturbance frequency and severity is rarely available at large scales and both are commonly described by statistical distributions. For instance, Johnson and Van Wagner (1985) used the exponential distribution to describe the occurrence probability of disturbances including wildfire, storm, and other natural disturbances. In this study, I characterize the probability density of the disturbance return interval in landscape as:

$$f(T;\lambda) = \begin{pmatrix} \frac{1}{\lambda} \cdot e^{\frac{-T}{\lambda}}, T > 0 \\ 0, T < 0 \end{pmatrix} \qquad \text{Eq. 4.4}$$

Correspondingly, the cumulative distribution function $F(T;\lambda)$is:

$$F(T;\lambda) = \begin{pmatrix} 1 - e^{\frac{-T}{\lambda}}, T > 0 \\ 0, T < 0 \end{pmatrix} \qquad \text{Eq. 4.5}$$

in both Eq. 4.4 and 4.5, T (yr) is the interval of two consecutive disturbance events, and λ (yr) is the MDI specific to the ecoregion or biome. In the landscape with a certain MDI, the probability that another disturbance occurs after the last disturbance increases as time goes by (Figure 4.1c).

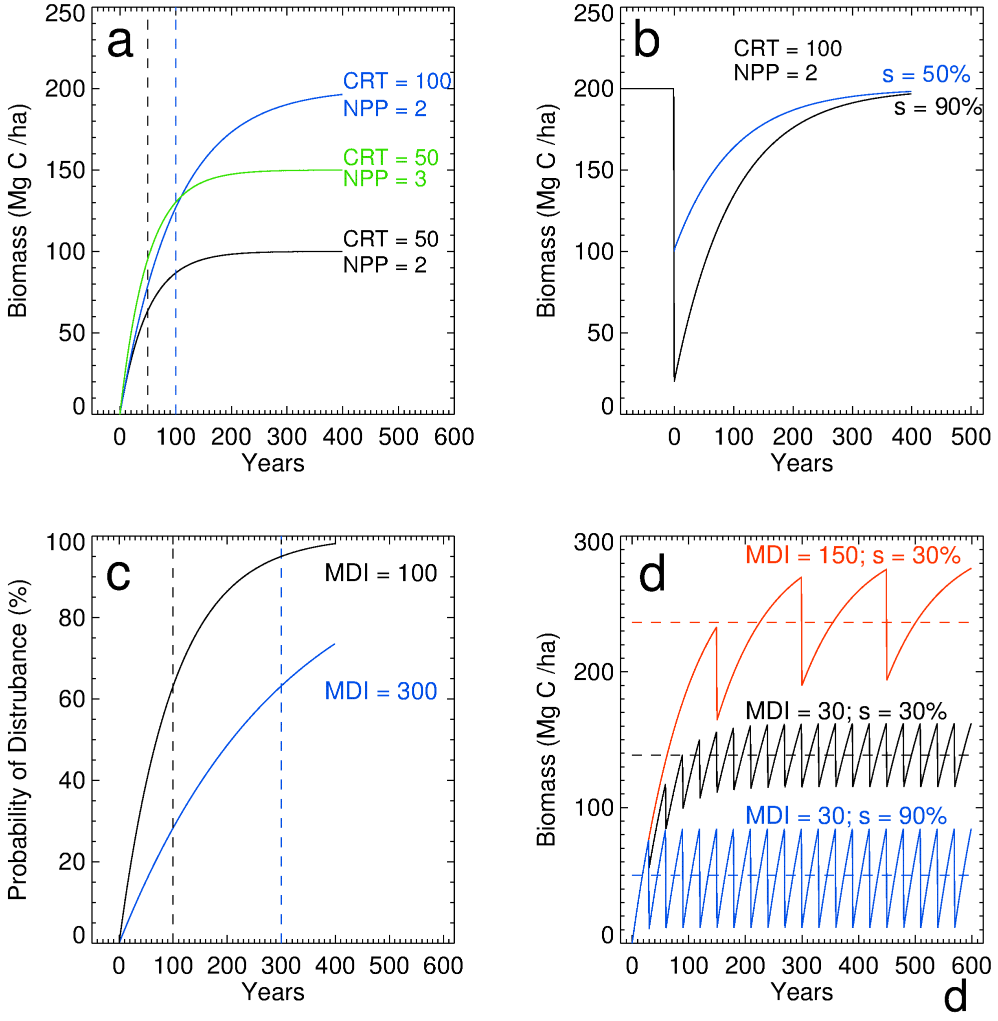


Figure 4.1: Model performance with varied parameters. (a) biomass (Eq. 4.1) simulated for variable NPP (Mg C/ha) and CRT (yr), dashed lines show CRT located at stand age axis. (b) biomass (Eq. 4.3) simulated for variable disturbance severity (s). (c) probability of disturbance (Eq. 4.5) simulated for variable MDI, dashed lines are MDI located at stand age axis. (d) biomass change under different disturbance regimes (MDI and s), dashed lines are the average biomass at quasi-equilibrium.

4.2.3 Solution of Disturbance Regime

By applying the recovery biomass accumulation equation (Eq. 4.3) in the exponential distribution of disturbance intervals (Eq. 4.5), I determine the expectation of the biomass (*E[X]*, gC/m^2) from disturbance severity (s) and the MDI (λ) with stable carbon influx in an equilibrium landscape (Figure 4.1d; Weng et al., in press),

$$E[B] = \frac{NPP \times \tau \times \lambda}{\lambda + s \times \tau} \qquad \text{Eq. 4.6}$$

Reversely the MDI (λ) is derived as,

$$\lambda = \frac{s \times NPP \times E[B]}{NPP \times \tau - E[B]} \qquad \text{Eq. 4.7}$$

4.2.4 Equilibrium Landscape Candidates

If the spatial scale of my analysis had been indefinite, equilibrium would no longer be relevant due to the distinctness of disturbance regimes and the discontinuities or gradients in topography, soils, moisture or other environmental factors. Shugart and West (1981) suggested that a quasi-steady state is likely to exist only at the landscapes much larger than the average size of the disturbances, for example, the landscape of a stable patch mosaic from 100,000 ha to 400,000 ha. Consequently, in this study I assessed the landscape at 0.5 degree resolution (about 300,000 ha), which contains thousands of pixels in 1 km resolution AGB and NPP map.

The scenario of landscape equilibrium ideally requires a primary forest inside each grid cell, because other vegetation types with distinctive disturbance behaviors may alter the theoretical construct outlined above. Landscape grids of the sparse forest cover mixed with grass, shrub or anthropogenic modification could introduce noise in the input data measurement and model simulation, like decreasing mean NPP and mean biomass. In this study, to avoid such noise in the representative equilibrium landscapes, all 0.5 degree grids must pass rigorous filtering: I only selected 0.5 degree grids with at least 90% forest cover. I calculated the forest fraction for each grid from the MODIS IGBP land cover product (MOD12Q1) at 500 m resolution in year 2005. Most of the candidate grids are within boreal and tropical forest; a few temperate forests passed the screening as well (Figure 4.2).

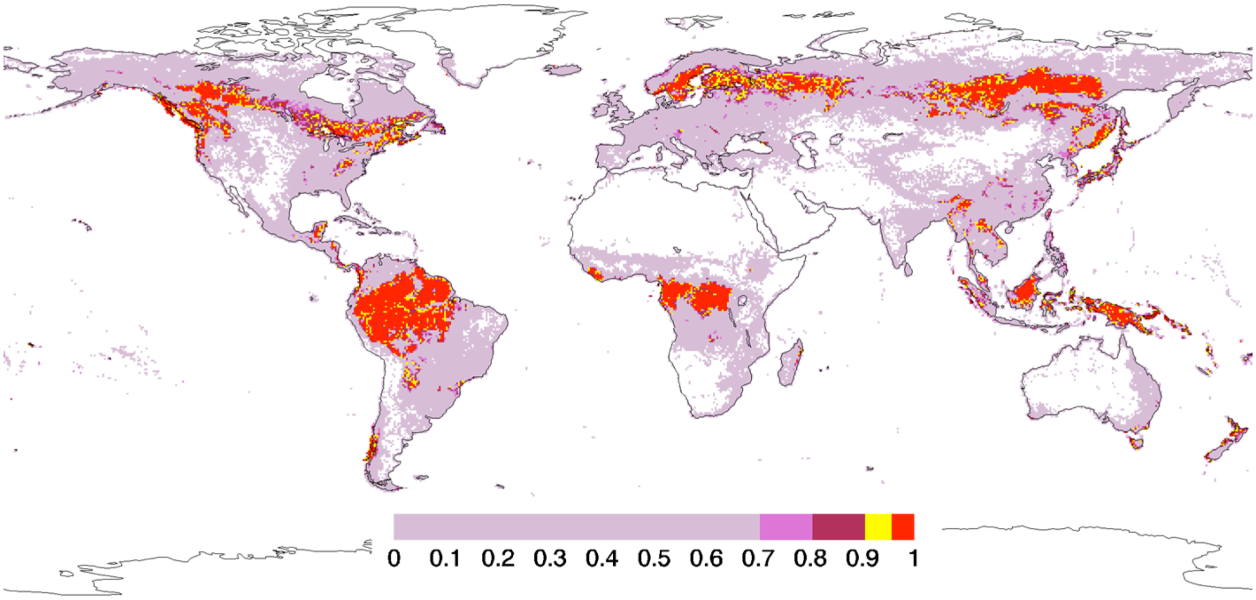


Figure 4.2: The percentage of global forest fraction in 0.5 degree grid. All forest types in MOD12Q1, including evergreen needle leaf forest, evergreen broadleaf forest, deciduous needleleaf forest, deciduous broadleaf forest, and mixed forest, were considered as forest. The forest fraction is the forest cover area divided by the land area for each grid.

4.2.5 Data

I assembled the data required to simulate stand age as described in the above equations and in Figure 4.3. Compared to the scarce measurements about belowground carbon pools, AGB is relatively easily estimated from available field observations and remote sensing (Houghton, 2005). Thus, I applied the disturbance model only for the live AGB in the equilibrium forest. To estimate the global disturbance regime, several independent data sets with global coverage are required at corresponding grids, including carbon influxes to forest AGB, current carbon storage in AGB, potential AGB, and turnover rate of carbon in AGB.

Passive sensors with visible and infrared bands are able to capture information about tree cover, leaf area, shadow and other canopy structure characteristics, while LiDAR or RADAR can penetrate into the canopy to obtain tree height and tree density. Combining passive and active multi-sensor data, Yu et al. (2010) developed a fusion algorithm for regional AGB monitoring, and validated the biomass with ground inventory. I used an updated 1 km global AGB map, based on a similar algorithm but driven by MODIS reflectances and LiDAR metrics (Saatchi et al., 2011).

I used 0.5 degree mean annual NPP derived from 8-day 1 km MODIS NPP data from 2001 to 2004 (MOD17A2). The NPP product has been validated with large footprint ground measurements (Turner et al., 2006), and, conceptually, is the annual total carbon influx under local climate conditions. Although NPP may rapidly accelerate in young forests and decrease with age (Luyssaert et al. 2008), preliminary sensitivity tests showed that the amount of potential carbon storage capacity was not significantly related to the short-term temporal change of NPP for the equilibrium forest (Weng et al., in press). Weng et al. (in press) found that biomass estimated with an assumption of constant carbon influx was negligibly different that biomass estimated with variable carbon inputs, at least when the residence period is longer than 20 years. In this study, therefore, I use the spatially averaged annual NPP in each 0.5 degree grid cell as the annual constant carbon influx.

The allocation of NPP to AGB and belowground biomass varies across different nutrient regimes and stand ages of forests (Gower et al., 2001). There are some general scaling rules established for NPP partitioning and estimates of aboveground NPP (ANPP). For example, Potter (1999) succeeded in running an ecosystem NPP model with empirical values of the partitioning rate. In this simple framework, I assumed half of total NPP is allocated to AGB (Potter, 1999).

Pan et al. (2011) used inventories from the United State Forest Inventory Analysis (FIA) and the Canadian national forest inventory (CanFI), fire data, remote sensing and an interpolation scheme to produce a 1km stand age map for North America. It provides the best if not the only detailed stand age for North America. I aggregate Pan's "interpolated map" to the 0.5 degree resolution of my “estimated map” and compared the two products. Finally, I conducted a comprehensive analysis of the effect of uncertainties in estimating grid parameters, remotely sensed NPP and AGB, and parameter values (see Auxilliary section for details).

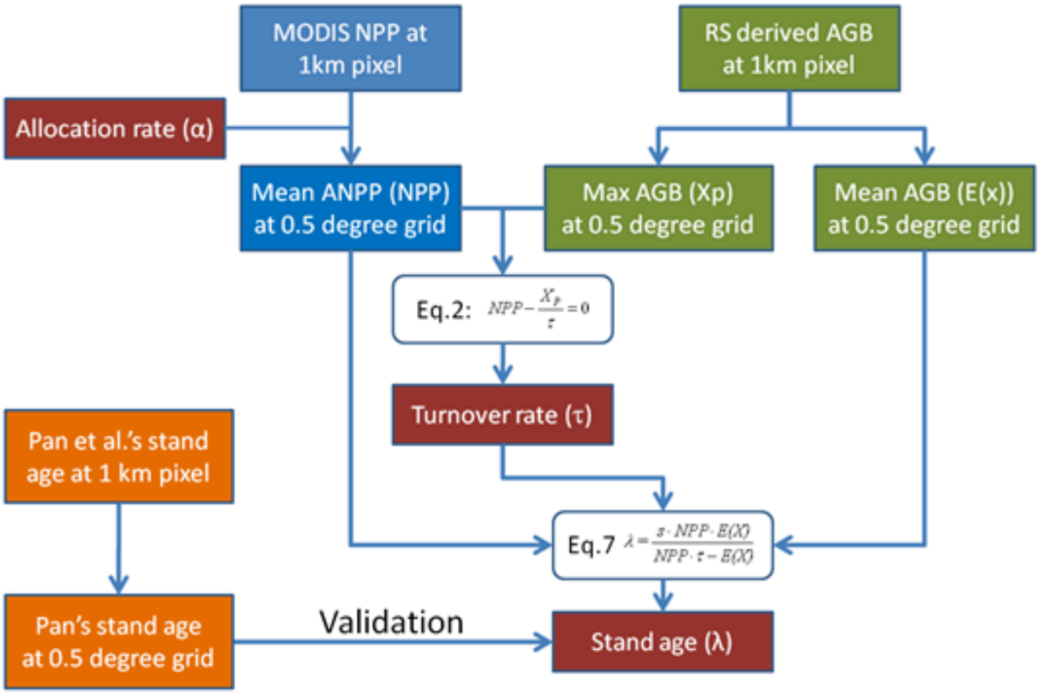


Figure 4.3: The framework of estimating turnover rate and stand age from remote sensing derived NPP and Biomass.

4.3 Results

4.3.1 Mean CRT Calculation

The mean CRT (calculated from maximum AGB and mean NPP, Eq. 4.2) of tropical forests is less than 50 years in most of the Amazon and Congo basins except some fringe region close to Atlantic Ocean, while CRT is between 50 and 100 years in Southeast Asia (Figure 4.4). In contrast, carbon in most boreal forests turns over at time scales longer than 100 years. The boreal forests in Eurasia with long residence time extend from Scandinavia to Siberia, though the

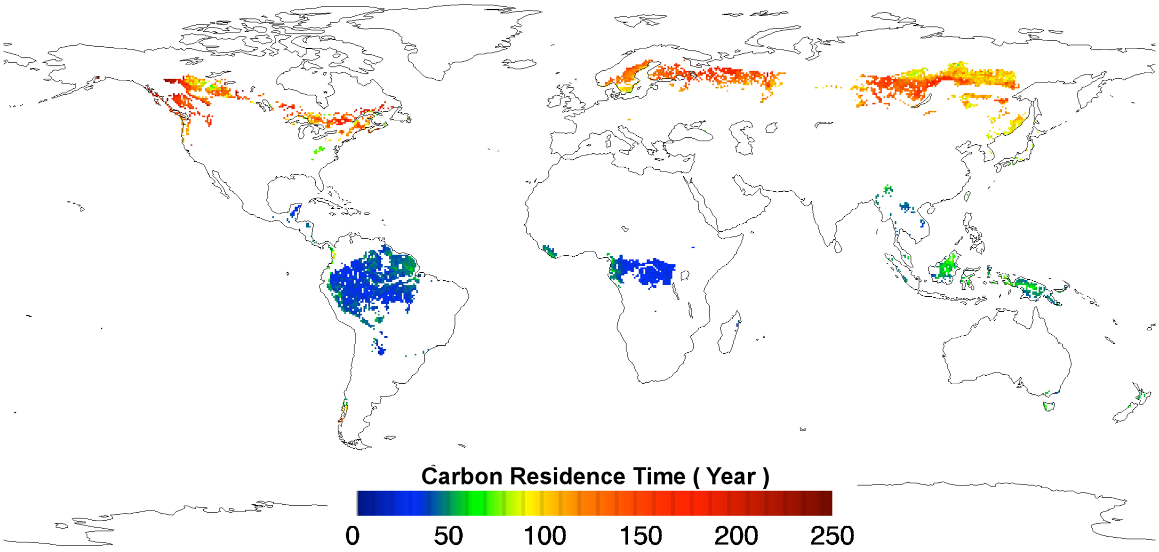


Figure 4.4: The CRT (year) for AGB of primary forest in 0.5 degree grids. The residence time is the reciprocal of mortality rate calculated from NPP and maximum AGB.

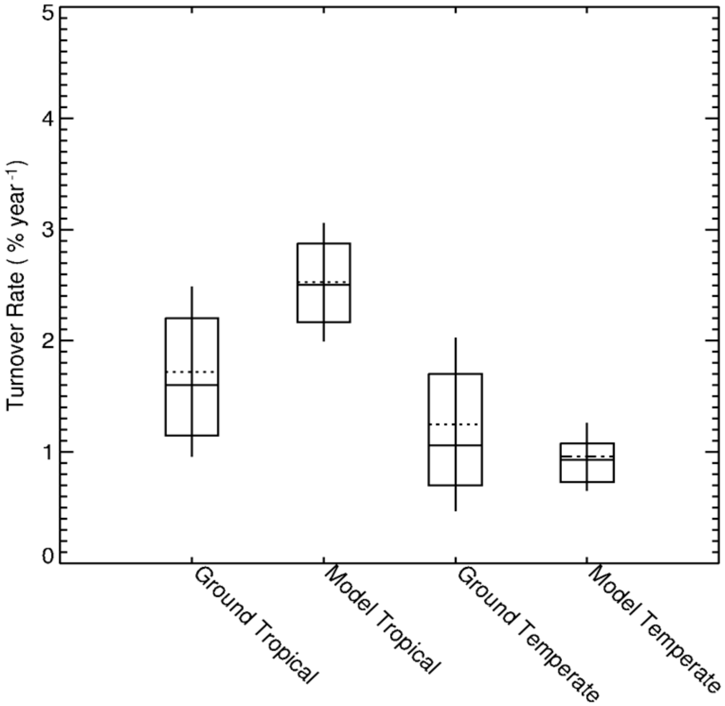


Figure 4.5: The observed (Stephenson and Mantgem, 2005) and estimated tropical and temperate forest mortality rates. Boxes are the standard deviation range. The vertical line encompasses the 25th through 75th percentiles; the solid horizontal lines indicate mean mortality rate. Dotted horizontal lines indicate the median.

CRT is slightly shorter than 100 years in northern Siberia, East Asia, and southern Scandinavia. Similarly, the CRT of boreal forests over North America is mostly over 100 years, consistent with those over Eurasia. The results above clearly show that the CRT of tropical forests is shorter than in boreal forests, qualitatively consistent with previous studies from ground measurement (e.g., Stephenson and Mantgem 2005, Figure 4.5).

4.3.2 Stand Age Estimation

I further estimated a stand age map for all equilibrium grids from NPP and live aboveground biomass based according to Eq. 4.7. As in the case of CRT, the results show that the stand ages for tropical forests are much shorter than temperate and boreal forests (Figure 4.6). Tropical forest MDI is usually lower than 50 years, but MDI in the northeastern Amazon, West Congo Basin, and New Guinea is between 50 and 100 years. In comparison, boreal forests in Eurasia exhibit high mean stand age from 100 to 300 years. Most core regions of boreal forests,

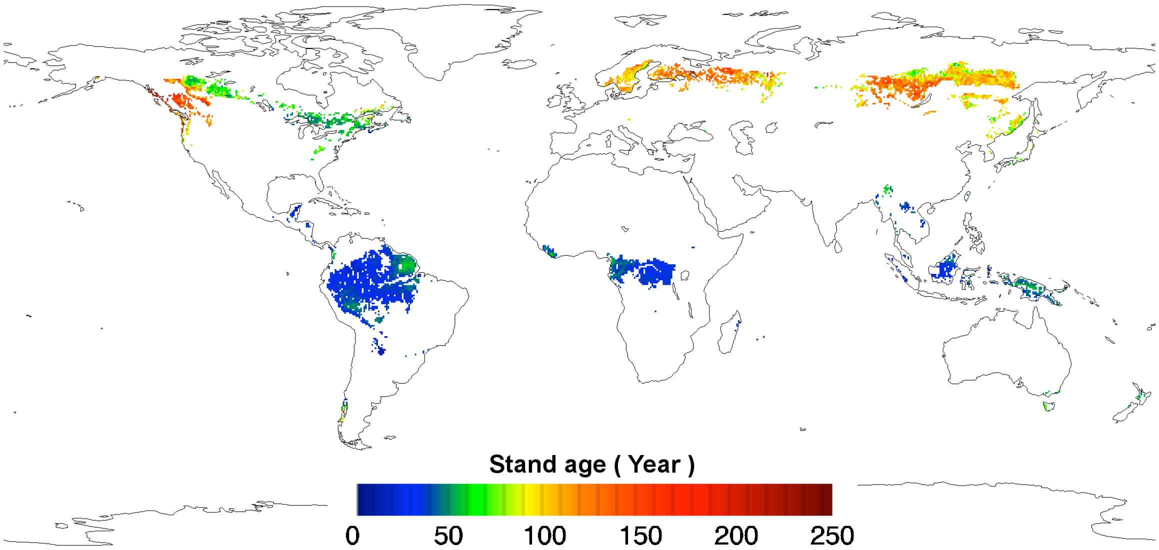


Figure 4.6: The mean stand age for all primary forest 0.5 degree pixels.

e.g. Western Russia and Central Siberia, are homogeneously covered with old growth forests. The forests in Eastern Siberia are of younger mean stand ages, where the CRT is short as well. Similarly, lower stand ages corresponding to short CRT are found in boreal forests along the coast of the Sea of Japan.

Unlike the spatial pattern of CRT, which is consistent over Eurasia, the estimated stand age of temperate and boreal forests displays a remarkable transition in North America. The estimated forest stand ages are about 150 years in the Pacific North-West regions, similar to the corresponding CRT. However, for the vast areas stretching from Canadian Rockies all the way to the East Coast (New England/Nova Scotia), the estimated forest stand ages are mostly ~50 years, much lower than the corresponding residence time (100-150 years).

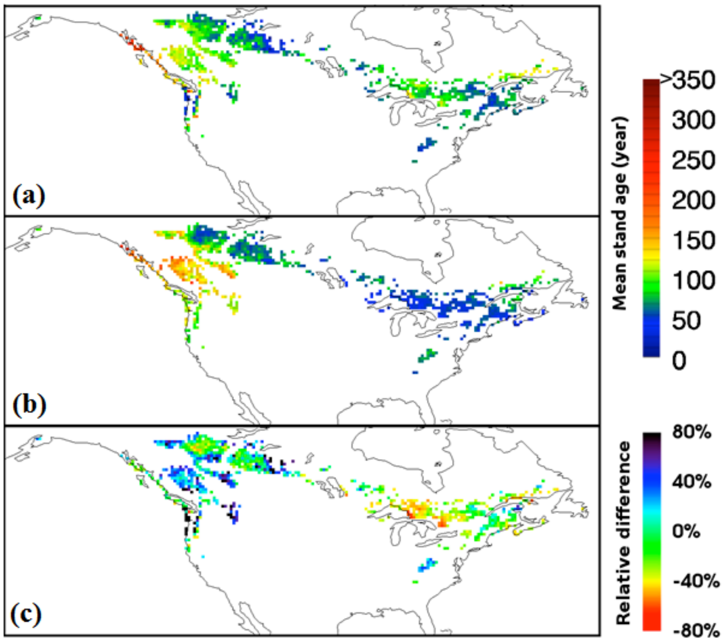


Figure 4.7: The mean stand age maps from Pan et al's interpolated method (a) and model estimation (b) on equilibrium grids. (c) is the relative difference between the two maps (b-a)/a*100%

The estimated mean stand age map of this study mostly agrees with the interpolated stand age map based on ground inventory data and recent disturbances (Figure 4.7). In both maps, the oldest forests are located in western Canada with 300 years of stand age followed by Montane Cordilleran forest older than 200 year. Both maps illustrate young forests from the Canadian taiga in middle regions to northern forests near the Atlantic Ocean. Regionally, the estimated map differs from the interpolated map in a few key respects: (1) higher values in the US west coast forest; (2) shorter MDIs for the northern forest grids near the Great Lakes; and (3) higher ages around Appalachian Mountains.

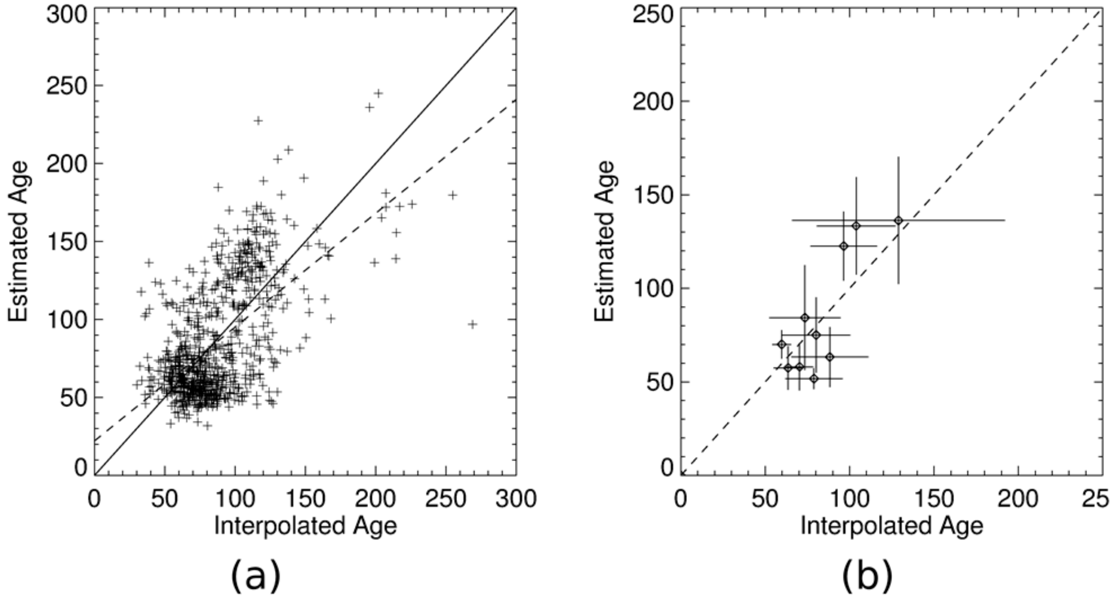


Figure 4.8: Comparison plots between estimated age and interpolated age. (a) The comparison between Pan's age map and the estimated age from my model (Unit: years). The correlation coefficient of all available 793 grids is 0.59, with bias at -1.25 year and root mean square deviation of 31.03 year. (b) Same comparison at ecoregional scale, the plots are aggregated to 10 major ecoregions. The diamonds are the mean stand age of each ecoregion, and the two bar crossing the diamond are the standard deviations of the two stand age maps inside each ecoregion.

I compared every pixel in the two maps (Figure 4.8a), which again shows that the estimated stand age generally agrees with interpolated map. There is a significant correlation (r =

0.58, $p<0.05$, n=802) between two maps within the whole range, and the bias of estimation is 1.3 year (Figure 4.4a). The estimated values scatter around the interpolated one, but the RMSE is modest (31 year) compared to the range of stand ages (0 to 400 years). The estimated stand age tends to be older in the middle range from 80 to 150 year, yet also underestimates the age for some very old forests. Based on the level II ecoregions of North America (Commission for Environmental Cooperation, 1997), I aggregated the two stand age maps into 10 ecoregions containing no less than 20 grids of 0.5 degree. The plots of ecoregional mean stand age from both maps are close to the 1:1 line, and the estimated stand age has a lower standard deviation in younger ecoregions (mean age < 90 years) than the interpolated map (Figure 4.8b).

4.3.3 Total Uncertainty

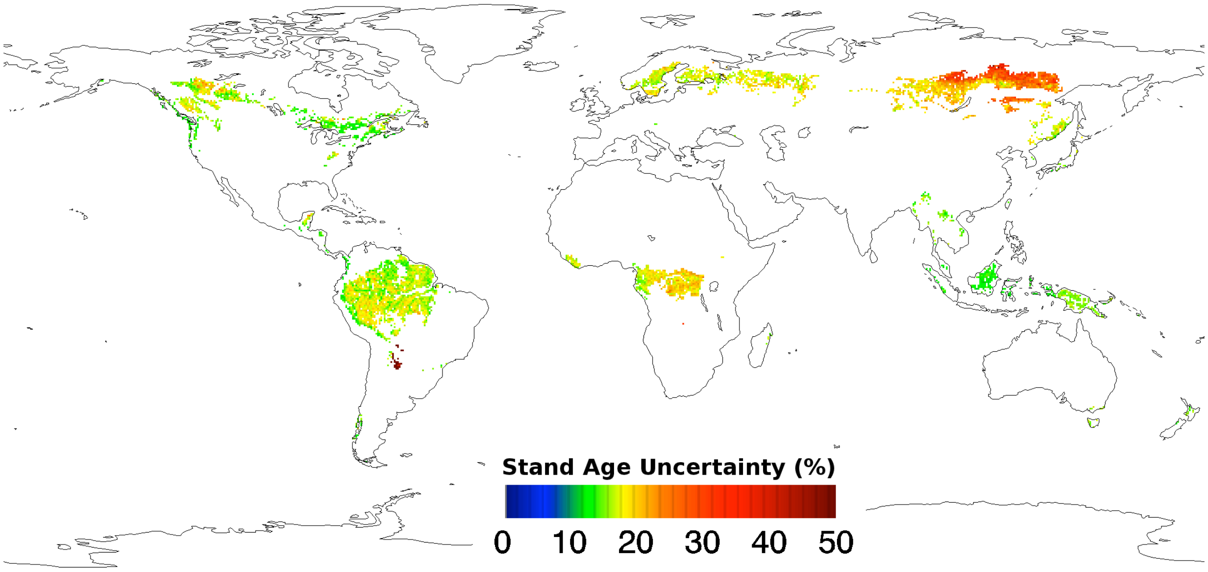


Figure 4.9: The uncertainty of estimated stand age by using Monte Carlo errors propagation method. The uncertainty is expressed as percentage of standard derivation against estimated value.

I calculated the total uncertainty of the stand age map from the uncertainties in grid parameters, remotely sensed NPP and AGB, and model parameters (see Auxilliary section, Figure

4.9). The eastern part of North America and Southeast Asia have lowest uncertainty (less than 15%), followed by western North America, European, Africa, and most of Amazonia with uncertainties from 10 to 25%. Siberia and small parts of South America have large uncertainties up to 50%, which may due to the small difference between mean and maximum AGB.

4.4 Discussion

This framework describes only natural disturbances of the primary forest equilibrium landscape, not anthropogenic disturbances which seldom follow the exponential distribution of disturbance probability (Eq. 4.4 and 4.5). The 90% forest land cover filter may eliminate regions with recent anthropogenic effects, but may not be able to exclude all human effects. Although over 35 percent of the global forest area is unmanaged primary forest according to FAO (2010), anthropogenic disturbance, nitrogen deposition, and other human activities extend the human footprint to the temperate and boreal forest (Magnani et al., 2007). For instance, forests in eastern North America have a short stand age, which is usually explained as the consequences of human development since the 19th Century (Pan et al., 2011). After decades of reforestation, it is difficult to filter the anthropogenic disturbance only by identifying the land cover. A human footprint map characterizing the anthropogenic disturbance separately from natural disturbances would help improve the results of simulation on primary forest. I used “The Last of the Wild” version 2 map (Sanderson et al., 2002) to select the ‘wild’ grid (that human footprint index no more than 10) from 0.5 degree candidates. The refined ‘wild’ grids led to an estimated stand age map closer to interpolated map (R= 0.68, Figure 4.10) but only half grids left.

Some particular ecoregions with detailed anthropogenic footprint mapping are now available (Peres, et al., 2006), and further refinements to my stand age framework could be implemented to remove human disturbed places, providing a potentially more accurate and practical guideline for programs like reduced emissions from deforestation and degradation

(REDD). On the other hand, the current stand age regime undergoing human effects, as on opposite of natural disturbance regime with similar climate background, provides a possible clue to track the historical anthropogenic disturbance, and to detect valuable carbon sequestration region. For example, Figure 4.4 indicates that boreal forests over North America are ~50 years younger than those over Eurasia. If anthropogenic disturbances are the main factor determining the short stand age of boreal forests in North America, it is implied that those forests may have a higher potential to sequester carbon than those in Eurasia. Therefore, I report the stand age map before human footprint filtering, which shows a global overview of disturbance with more spatial coverage, and could hint the region with anthropogenic disturbances.

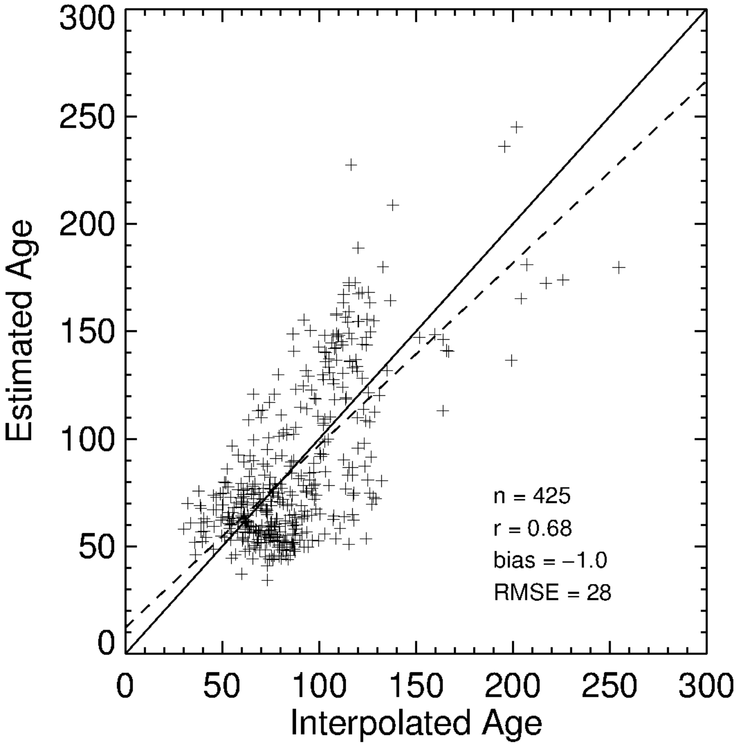


Figure 4.10 Comparison of 0.5 degree grids filtered by human footprint map (Sanderson et al. 2002) against interpolated map in North America

Moreover, the assumption on equilibrium landscape also requires steady states of carbon storage and fluxes. But some recent and extreme disturbances with impacts over vast areas may lead to unconvincing disturbance regime. It is beneficial to eliminate the regions which suffered from recent extremely disturbances by using the remote sensing derived records about burnt regions during last decades (Barbosa et al., 1999; Roy et al., 2008).

4.5 Conclusion

A mechanistic framework for generating stand age of equilibrium forest based on several basic assumptions was explained in this study, and it revealed a theoretical pathway linking historical disturbance with independent current AGB and NPP measurement. The mean stand age map produced in this study is the first analysis of all primary forests globally based on remote sensing data. It provides a fresh view of understanding long-term ecosystem dynamics and states from short-term remote sensing data. The model was validated with the interpolated age map from ground inventories in North America, and the estimated age map generally agrees with the interpolated map in spatial pattern and magnitude of stand age in both pixel and ecoregion scale, with some differences in particular regions. The analysis on assumptions addressed the limitation usage of this model, especially for tropic forest age estimation.

4.6 Uncertainty Analysis

Since the stand age estimation integrated remote sensing and theoretical ecosystem models under some key assumption, it is necessary to assess the uncertainty of the turnover and stand age maps. The uncertainty analysis of the whole framework involved: (1) statistical errors in representing biomass distribution inside equilibrium grids, including landcover, grid size, maximum and mean biomass values; (2) observation errors of the remote sensing derived NPP and AGB; (3) prediction errors of parameters in the framework formulation, such as aboveground

NPP allocation rate and mean disturbance severity.

4.6.1 Statistical Errors in Biomass Distribution

In this study, I used the maximum biomass to represent the potential biomass inside an equilibrium grid. To test the sensitivity of the maximum biomass estimate, I used the 85th, 90th, and 95thpercentiles of sub-grid AGB, instead of maximum value, to calculate CRT. The results were not significantly different from each other concerning the residence time (turnover rate increased less than 5%). On the other hand, it is possible that the size of the grid may still not be large enough to represent the regional equilibrium, and the grid maximum value may correspondingly be smaller than the true potential maximum. If the grid size is too large, however, only a few candidates without anthropogenic effect would meet the natural equilibrium scenario. The dilemma between the spatial scale and assumption of natural equilibrium poses an obstacle to this study and remains to be solved in the future.

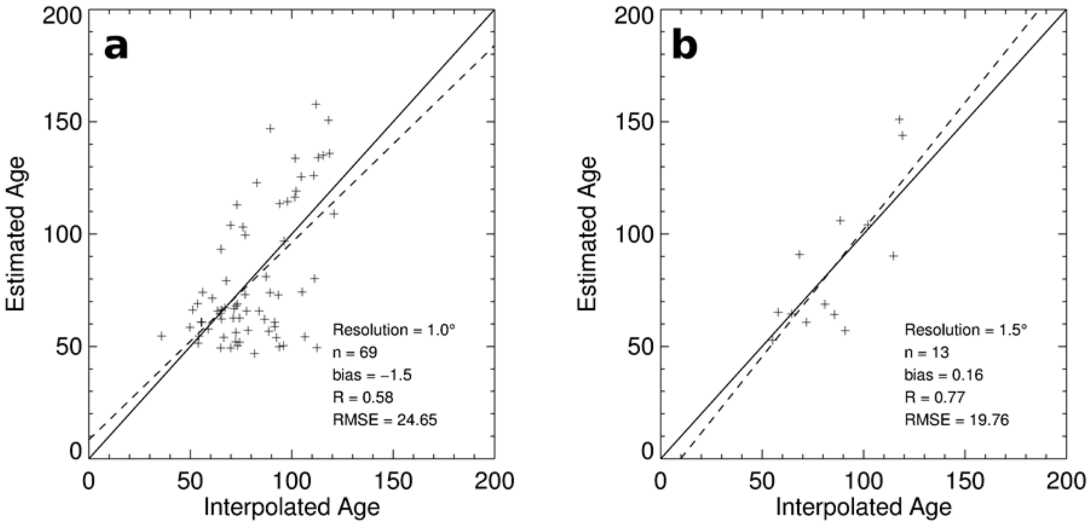


Figure 4.11: Grid sensitive plots: comparison of predicted stand age calculated at a 1 degree (a) and 1.5 degree (b) resolutions against interpolated map in North America.

I next executed the framework at 1 and 1.5 degree landscape grids to exam the framework performance along spatial scale as well (Figure 4.11a&b). Compared to Pan's interpolated age map, the estimation at 1 degree is similar to original 0.5 degree results, while results on 1.5 degree grid shows a higher correlation but with fewer grids. This suggests that the methods works better for a larger grid with more biomass samples, if noises from land cover and human effect are filtered.

I also assumed that the current mean biomass in a grid cell is the true biomass of an equilibrium forest, but in reality, it could represent the biomass of recovering or disturbed forest. Recent research (Iwao et al., 2011) shows that the MODIS landcover product has 78.2% agreement with a vast field data for global forest land.

4.6.2 Observation Errors of the Remote Sensing Derived NPP and AGB

The four year average MODIS NPP data employed in this study may not represent the average status of carbon fluxes, considering the long-term trend of carbon fluxes influenced by climate change, such as global NPP decreasing (Zhao and Running, 2010) and mortality increasing in Canada's boreal forests (Peng et al, 2011). But those are difficult to involve in this uncertainty analysis.

The MODIS NPP production was validated with large footprint ground data (25km2), and gave the ratio of root mean square error to the mean of the measured annual NPP at average 0.26 (Turner et al. 2006). Since I used the 0.5 degree average NPP, the grid error should be

$$\varepsilon NPP_{grid} = \frac{\sqrt{\sum_{i=1}^{N}(NPP_i \varepsilon NPP_i)^2}}{\sum_{i=1}^{N} NPP_i} \qquad \text{Eq. 4.8}$$

where, N is the number of pixels in grid, NPP_i is the pixel level NPP, εNPP_i is the error at pixel level, and εNPP_{grid} is the error at grid level. Since N is getting large when grid size increased, the relative error will quickly decline. Hence, the error of grid average NPP is much lower than pixel level, for example the average NPP error stay below 3% for N=50 (1km pixels in 0.5 degree). So the average AGB value at 0.5 degree also has low errors range around 2% compared to over 20% error at 1km pixel (Saatchi et al. 2011). Since both the maximum biomass and 90th percentile biomass lead to similar result of turnover, the maximum biomass should represent the top 10 percent biomass pixels in grid (N=25). So I set the maximum AGB error as 4%.

4.6.3 Prediction Errors of Parameters in Formulation

The NPP partitioning to aboveground –ANPP– is set as half of total NPP to simplify the framework. But this ratio is different for various vegetations, and also varies along the growth period for same vegetation. Although Litton et al.(2007)'s review shown small changes of partitioning in response to age, the spatial variability of ANPP to NPP ratios are report from 0.35 to 0.50 (Luyssaert 2007). The relative high ANPP based on 0.5 partitioning rate may lead to underestimation of CRT and stand age based on Eq 2&7.

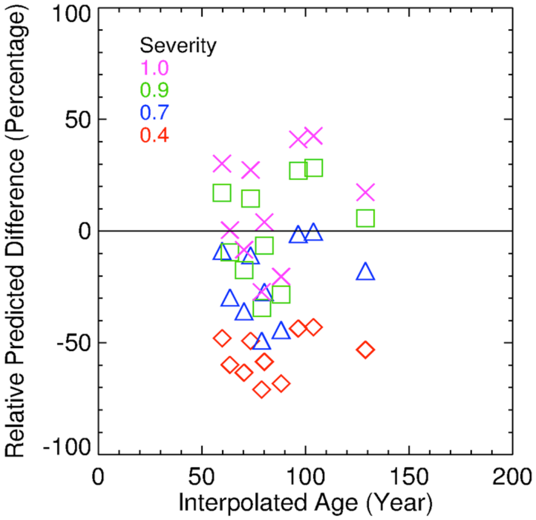


Figure 4.12 The relative differences between four estimated mean stand age and mean interpolated stand age (Pan et al 2011). The mean stand ages are estimated based on four severity scenario, 0.4 in red, 0.7 in blue, 0.9 in green, and 1.0 in pink.

In estimating the CRT, I have assumed the impacts of natural disturbances on vegetation to be severe (>=90% biomass are removed), so the varied status of forest biomass during the recovery period can be clearly detected by fitting the distribution of disturbance. This assumption is particularly reasonable for high-latitude primary forests, which are mostly disturbed by wildfire with significantly clear-cut burned area over 1000ha (Bergeron et al., 2002). The sensitivity of disturbance severity is explored by setting model with the severity at 0.5, 0.75 and 1.0. Compared to interpolated age map, the stand ages estimated with low severity are generally younger than the interpolated age, and high severity results less bias in North America (Figure 4.12). It also implied the high severity assumption is suitable for high-latitude region.

However, various nature disturbances in tropic forests may not have as high severity as the wildfire does in boreal forests (e.g. Espírito‐Santo et al. 2010), and the severity of disturbance are related to vegetation type and tree density (Frelich and Reich 1999). As such, forests within a grid cell may not be completely cut-off but with much legacy biomass, which does not match the high severity parameter in the model. Figure 4.8 illustrated the limitation of low severity and fast turnover on the AGB-to-MDI pathway. The difference of MDIs in high severity and slow mortality scenario will lead to large variance of average biomass like the wildfire in boreal forest (Figure 4.13a), but the biomass values with same MDI gradient are not distinct for low severity and fast turnover rate scenario (Figure 4.13b & Figure 4.13c).

In this study, the estimations of CRT time and stand age based on the framework described in Eq. 4.3 and Eq. 4.7 may not be suitable for tropic forests. The mixture of disturbance and growth within one pixel weakens the no-disturbance assumption in CRT estimation. Hence the turnover rate for tropic forest may be overestimated by coupling inside-pixel-cutting effects. The biomass of tropical forests after low severity disturbance restores quickly due to high NPP, and thus there are not enough pixels at recovery period with below average biomass to determine the parameter of disturbance distribution. Although there is limit knowledge about severity value for global vegetation, the uncertainty of severity on stand age is due to the linear relation between severity and MDI. Here the possible severity range is set from 0.6 to 1.0.

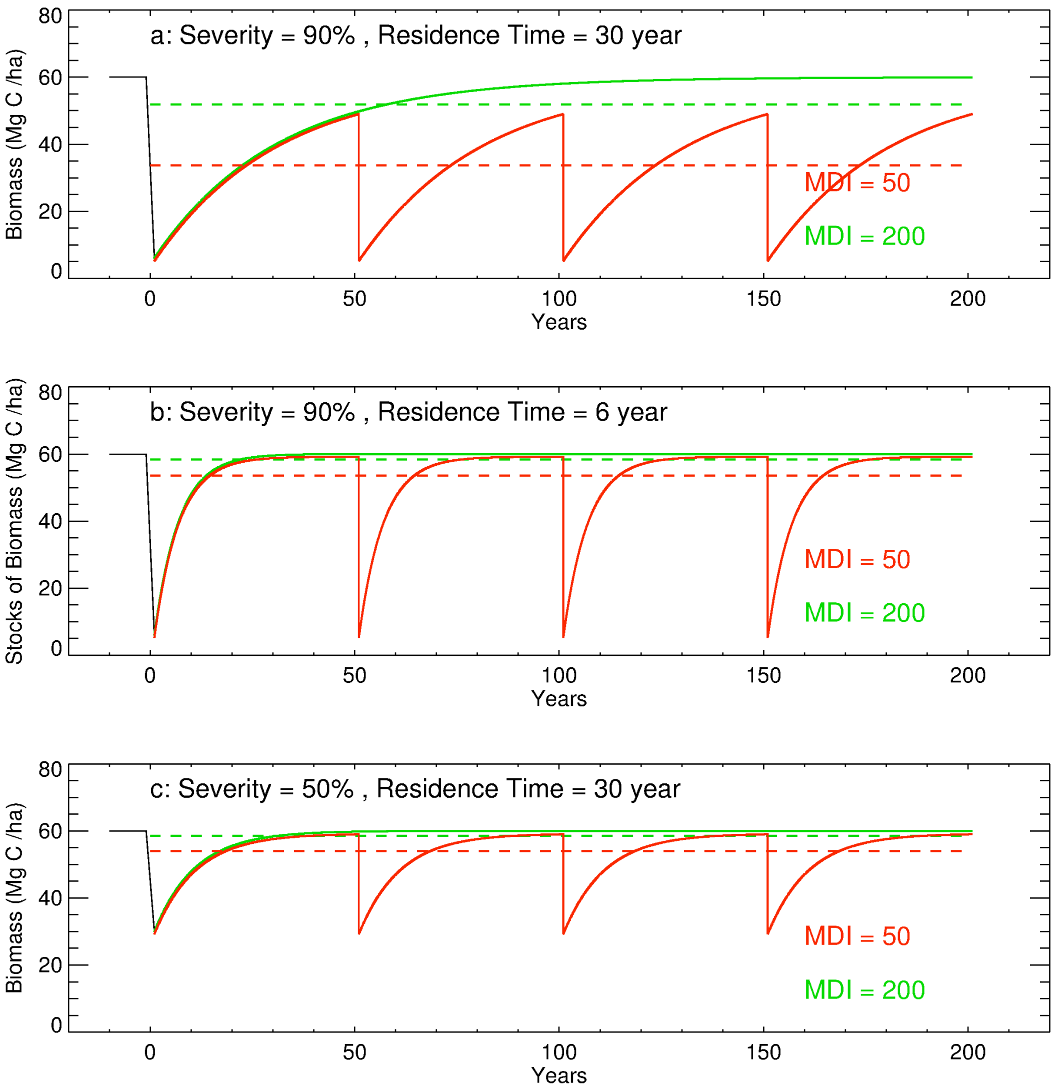


Figure 4.13 Scenarios of disturbance severity and mortality rate. (a) severity =80%, mortality = 30 year; (b) severity = 80%, mortality = 6 year; (c) severity = 50%, mortality = 30 year. The average biomass (dashed line) with three disturbance regimes: red (MDI) = 66 year); blue (MDI = 100 year); green (MDI = 200) are shown in each scenario.

4.6.4 Total Uncertainty Analysis

With above input variable error range (maximum AGB at 4%, mean ABG at 3%, mean NPP at 2%, partitioning ratio from 0.35 to 0.50, severity from 0.8 to 1.0), I built a Monte Carlo approach- making random 2000 values inside the error range for each input variable and parameter separately, then ran the model with 2000 input variable and parameter vector and produced the output vector. Then I calculated the standard derivation of output vector as total uncertainty in stand age estimation.

References

Achard, F., Eva, H.D., Stibig, H.J., Mayaux, P., Gallego, J., Richards, T., Malingreau, J.P., 2002. Determination of deforestation rates of the world's humid tropical forests. *Science* 297, 999.

Baldocchi, D., 2008. TURNER REVIEW No. 15.'Breathing'of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. *Australian Journal of Botany* 56, 1–26.

Barbosa, P.M., Grégoire, J.M., Pereira, J.M.., 1999. An algorithm for extracting burned areas from time series of AVHRR GAC data applied at a continental scale. *Remote Sensing of Environment* 69, 253–263.

Bender, M.A., Knutson, T.R., Tuleya, R.E., Sirutis, J.J., Vecchi, G.A., Garner, S.T., Held, I.M., 2010. Modeled impact of anthropogenic warming on the frequency of intense Atlantic hurricanes. *Science* 327, 454–458.

Bergeron, Y., Leduc, A., Harvey, B.D., Gauthier, S., 2002. Natural fire regime: a guide for sustainable management of the Canadian boreal forest. *Silva Fennica* 36, 81–95.

Canadell, J.G., Raupach, M.R., 2008. Managing forests for climate change mitigation. *Science* 320, 1456–1457.

Chambers, J.Q., Fisher, J.I., Zeng, H., Chapman, E.L., Baker, D.B., Hurtt, G.C., 2007. Hurricane Katrina's carbon footprint on US Gulf Coast forests. *Science* 318, 1107–1107.

Commission for Environmental Cooperation. 1997. Ecological regions of North America, Level II. (http://www.cec.org/naatlas/)

Chen, J.M., Ju, W., Cihlar, J., Price, D., Liu, J., Chen, W., et al. 2003. Spatial distribution of carbon sources and sinks in Canada's forests. *Tellus B* 55, 622–641.

Connell, J.H., Sousa, W.P., 1983. On the evidence needed to judge ecological stability or persistence. *American Naturalist* 789–824.

DeAngelis, D.L., Waterhouse, J.C., 1987. Equilibrium and nonequilibrium concepts in ecological models. *Ecological monographs* 1–21.

Espírito-Santo, F.D.B., Keller, M., Braswell, B., Nelson, B.W., Frolking, S., Vicente, G., 2010. Storm intensity and old-growth forest disturbances in the Amazon region. *Geophysical Research Letters 37*, L11403.

Fahey, T.J., Woodbury, P.B., Battles, J.J., Goodale, C.L., Hamburg, S.P., Ollinger, S.V., Woodall, C.W., 2009. Forest carbon storage: ecology, management, and policy. *Frontiers in Ecology and the Environment* 8, 245–252.

Franklin, J.F., Spies, T.A., Pelt, R.V., Carey, A.B., Thornburgh, D.A., Berg, D.R., Lindenmayer, D.B., Harmon, M.E., Keeton, W.S., Shaw, D.C., 2002. Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. *Forest Ecology and Management* 155, 399–423.

FAO (United Nations Food and Agriculture Organization). 2010. Global Forest Resources Assessment 2010. Online at http://www.fao.org/forestry/fra/en/. Country Reports can be found at http://www.fao.org/forestry/fra/67090/en/.

Frelich, L.E., Reich, P.B., 1999. Minireviews: neighborhood effects, disturbance severity, and community stability in forests. *Ecosystems* 2, 151–166.

Goetz, S., Sun, M., Baccini, A., Beck, P., 2010. Synergistic use of spaceborne lidar and optical imagery for assessing forest disturbance: An Alaska case study. *Journal of Geophysical Research* 115, G00E07.

Gower, S.T., Krankina, O., Olson, R.J., Apps, M., Linder, S., Wang, C., 2001. NET PRIMARY PRODUCTION AND CARBON ALLOCATION PATTERNS OF BOREAL FOREST ECOSYSTEMS. *Ecological Applications* 11, 1395–1411.

Gurney, K.R., Raymond, L., 2008. Targeting deforestation rates in climate change policy: a “Preservation Pathway” approach. *Carbon Balance Manag* 3, 2.

Houghton, R.A., 2005. Aboveground Forest Biomass and the Global Carbon Balance. *Global Change Biology* 11, 945–958.

Huston, M.A., 2004. Management strategies for plant invasions: manipulating productivity, disturbance, and competition. *Diversity and Distributions* 10, 167–178.

Ichoku, C., Kaufman, Y.J., Giglio, L., Li, Z., Fraser, R.H., Jin, J.Z., Park, W.M., 2003. Comparative analysis of daytime fire detection algorithms using AVHRR data for the 1995 fire season in Canada: Perspective for MODIS. *International journal of remote sensing* 24, 1669–1690.

Iwao, K., Nasahara, K.N., Kinoshita, T., Yamagata, Y., Patton, D., Tsuchida, S., 2011. Creation of New Global Land Cover Map with Map Integration. *Journal of Geographic Information System* 3, 160–165.

Keith, H., Mackey, B., Berry, S., Lindenmayer, D., Gibbons, P., 2010. Estimating carbon carrying capacity in natural forest ecosystems across heterogeneous landscapes: addressing sources of error. *Global Change Biology* 16, 2971–2989.

Keith, H., Mackey, B.G., Lindenmayer, D.B., 2009. Re-evaluation of forest biomass carbon stocks and lessons from the world's most carbon-dense forests. *Proceedings of the National Academy of Sciences* 106, 11635–11640.

Kurz, W.A., Dymond, C.C., Stinson, G., Rampley, G.J., Neilson, E.T., Carroll, A.L., Ebata, T., Safranyik, L., 2008. Mountain pine beetle and forest carbon feedback to climate change. *Nature* 452, 987–990.

Lentile, L.B., Holden, Z.A., Smith, A.M.., Falkowski, M.J., Hudak, A.T., Morgan, P., Lewis, S.A., Gessler, P.E., Benson, N.C., 2006. Remote sensing techniques to assess active fire characteristics and post-fire effects. *International Journal of Wildland Fire* 15, 319–345.

Litton, C.M., Raich, J.W., Ryan, M.G., 2007. Carbon allocation in forest ecosystems. *Global Change Biolog*y 13, 2089–2109.

Luo, Y., Weng, E., 2010. Dynamic disequilibrium of the terrestrial carbon cycle under global change. *Trends in Ecology & Evolution*.

Luo, Y., White, L.W., Canadell, J.G., DeLucia, E.H., Ellsworth, D.S., Finzi, A., Lichter, J., Schlesinger, W.H., 2003. Sustainability of terrestrial carbon sequestration: a case study in Duke Forest with inversion approach. *Global Biogeochemical Cycles* 17, 1021.

Luyssaert, S., Schulze, E.-D., Borner, A., Knohl, A., Hessenmoller, D., Law, B.E., Ciais, P., Grace, J., 2008. Old-growth forests as global carbon sinks. *Nature* 455, 213–215.

Magnani, F., Mencuccini, M., Borghetti, M., Berbigier, P., Berninger, F., Delzon, S., Grelle, A., Hari, P., Jarvis, P.G., Kolari, P., Kowalski, A.S., Lankreijer, H., Law, B.E., Lindroth, A., Loustau, D., Manca, G., Moncrieff, J.B., Rayment, M., Tedeschi, V., Valentini, R., Grace, J., 2007. The human footprint in the carbon cycle of temperate and boreal forests. *Nature* 447, 849–851.

Mildrexler, D.J., Zhao, M., Heinsch, F.A., Running, S.W., 2007. A new satellite-based methodology for continental-scale disturbance detection. *Ecological Applications* 17, 235–250.

O'Halloran, T.L., Law, B.E., Goulden, M.L., Wang, Z., Barr, J.G., Schaaf, C., Brown, M., Fuentes, J.D., Göckede, M., Black, A., 2011. Radiative forcing of natural forest disturbances. *Global Change Biology*.

Oliver, C.D., Larson, B.C., Oliver, C.D., 1990. *Forest stand dynamics*. McGraw-Hill New York.

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., 2011. A large and persistent carbon sink in the world's forests. *Science* 333, 988–993.

Parton, W.J., Scurlock, J.M.O., Ojima, D.S., Gilmanov, T.G., Scholes, R.J., Schimel, D.S., Kirchner, T., Menaut, J.C., Seastedt, T., Moya, E.G., 1993. Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide. *Global biogeochemical cycles* 7, 785–809.

Peng, C., Ma, Z., Lei, X., Zhu, Q., Chen, H., Wang, W., Liu, S., Li, W., Fang, X., Zhou, X., 2011. A drought-induced pervasive increase in tree mortality across Canada's boreal forests. *Nature Climate Change*.

Peres, C.A., Barlow, J., Laurance, W.F., 2006. Detecting anthropogenic disturbance in tropical forests. *Trends in Ecology & Evolution* 21, 227–229.

Potter, C.S., 1999. Terrestrial biomass and the effects of deforestation on the global carbon cycle. *BIOSCIENCE-WASHINGTON-* 49, 769–780.

Roy, D.P., Boschetti, L., Justice, C.O., Ju, J., 2008. The collection 5 MODIS burned area product--Global evaluation by comparison with the MODIS active fire product. *Remote Sensing of Environment* 112, 3690–3707.

Running, S.W., 2008. Ecosystem disturbance, carbon, and climate. *Science* 321, 652–653.

Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M., Mitchard, E.T.A., Salas, W., Zutta, B.R., Buermann, W., Lewis, S.L., Hagen, S., Petrova, S., White, L., Silman, M., Morel, A., 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences* 108, 9899 –9904.

Sanderson, E.W., Jaiteh, M., Levy, M.A., Redford, K.H., Wannebo, A.V., Woolmer, G., 2002. The human footprint and the last of the wild. *BioScience* 52, 891–904.

Shugart, H.H., West, D.C., 1981. Long-Term Dynamics of Forest Ecosystems: Computer simulation models, which allow for numerous seedlings and the long lives of large trees, predict how forests will respond to different management techniques. *American Scientist* 69, 647–652.

Smithwick, E.A.., Harmon, M.E., Domingo, J.B., 2007. Changing temporal patterns of forest carbon stores and net ecosystem carbon balance: the stand to landscape transformation. *Landscape Ecology* 22, 77–94.

Stephenson, N.L., Mantgem, P.J., 2005. Forest turnover rates follow global and regional patterns of productivity. *Ecology Letters* 8, 524–531.

Sun, G., Ranson, K.J., Kimes, D.S., Blair, J.B., Kovacs, K., 2008. Forest vertical structure from GLAS: An evaluation using LVIS and SRTM data. *Remote Sensing of Environment* 112, 107–117.

Tilman, D., 1985. The resource-ratio hypothesis of plant succession. *American Naturalist* 827–852.

Turetsky, M.R., Kane, E.S., Harden, J.W., Ottmar, R.D., Manies, K.L., Hoy, E., Kasischke, E.S., 2010. Recent acceleration of biomass burning and carbon losses in Alaskan forests and peatlands. *Nature Geoscience* 4, 27–31.

Turner, D.P., Ritts, W.D., Cohen, W.B., Gower, S.T., Running, S.W., Zhao, M., Costa, M.H., Kirschbaum, A.A., Ham, J.M., Saleska, S.R., 2006. Evaluation of MODIS NPP and GPP products across multiple biomes. *Remote Sensing of Environment* 102, 282–292.

Turner, M.G., 2010. Disturbance and landscape dynamics in a changing world 1. *Ecology* 91, 2833–2849.

Turner, M.G., Romme, W.H., Gardner, R.H., O'Neill, R.V., Kratz, T.K., 1993. A revised concept of landscape equilibrium: disturbance and stability on scaled landscapes. *Landscape Ecology* 8, 213–227.

Van Wagner, C.E., 1978. Age-class distribution and the forest fire cycle. *Canadian Journal of Forest Research* 8, 220–227.

Weir, J.M.H., Johnson, E.A., Miyanishi, K., 2000. Fire frequency and the spatial age mosaic of the mixed-wood boreal forest in western Canada. *Ecological Applications* 10, 1162–1177.

Weng, E., Luo, Y., Wang, W., Wang, H., Hayes, D., McGuire, D., Hastings, A., Schimel, D., in press. Ecosystem carbon storage capacity as affected by disturbance regimes: a general theoretical model. *Biogeosciences*

Westerling, A.L., Turner, M.G., Smithwick, E.A.., Romme, W.H., Ryan, M.G., 2011. Continued warming could transform Greater Yellowstone fire regimes by mid-21st century. *Proceedings of the National Academy of Sciences* 108, 13165–13170.

Yu, Y., Saatchi, S., Heath, L.S., LaPoint, E., Myneni, R., Knyazikhin, Y., 2010. Regional distribution of forest height and biomass from multisensor data fusion. *Journal of. Geophysical. Research* 115, G00E12.

Zhao, M., Running, S.W., 2010. Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science* 329, 940.

CHAPTER 5

CONCLUSION AND FUTURE WORK

The conclusion of this dissertation is some key variables either short-term or long-term could be extracted from remote sensing data by integrating ecosystem models, as it has been shown in estimation of phenology, biomass, and stand age.

5.1 Contribution

The overall contribution of the research reported in this dissertation is an exploration of the application of remote sensing data to the problem of specifying and verifying the structure, function, and behavior of terrestrial ecosystem, and a demonstration that such applications of remote sensing data are both feasible and portable in terrestrial vegetation research, especially for the region without extensive ground measurement of vegetation. The contributions include:

- Validating the MODIS phenology products with ground vegetation phenology data and other existing remote sensing phenology indexes.
- Standardizing the plant phenology events of various species from different networks as uniform index representing vegetation phenology of local plant community.
- Exploiting the linear relationship between Landsat-based Leaf Area Index (LAI) and canopy height metrics derived from the ICESat GLAS measurements in California at 30 m resolution, and then deriving high-resolution (30-m) spatially continuous estimates of canopy height and AGB.
- Establishing the standard comparison theme against various AGB density map at different resolution by aggregating the density map to regional total biomass.

- Acquiring the first stand age map for global primary forests from the remote sensing derived NPP and AGB data based on theoretical framework between carbon flux, carbon storage, and disturbance regime.

5.2 Future Research Directions

Even though the framework of three topics has been greatly extended by the work described in this dissertation, there are still plenty of research opportunities.

- The ground phenology dataset need including recent ground observation after 2006, such national phenology network with standard protocol.
- People are more interested in the change of vegetation, but the interannual trend analysis of MODIS phenology products is absent in this study due to the limit data from 2001 to 2006. The remote sensing derived phenology could reflect the trend or change of vegetation phenology, since it may have significant but consistent bias compared to ground observations. It would be valuable to extend this research to the latest dataset and to explore the interannual change of MODIS phenology data.
- I have shown that AGB density decreases at coarser spatial resolution. To overcome this problem, I plan to build an AGB density at sample regions by aggregating tree-level AGB at a desired grid size, which is the best validation data for the satellite AGB density map
- This study address the possibility of theoretical disturbance model applied in particular region, which is curial for recent REDD and carbon credit system. With a detailed land cover map and specific parameters (such as NPP allocation rate and mean severity of local disturbance), it is very likely to produce a improved stand age map and carbon carry capacity estimates.

- Besides the disturbance, other environmental factors, such soil, slope, and water with spatially heterogeneity, may also lead to the biomass mosaic inside equilibrium forest grid. It is valuable to refine the framework on some regions with detailed soil, topography and hydrology map.

CURRICULUM VITAE

Gong Zhang

zhanggong07@gmail.com
435-890-3925

Education:

8/2007 - 7/2012 (expected)	**Ph.D in Ecology**, Utah State University, Logan, UT
9/2004 - 6/2007	**MS in Geography**, Peking University, Beijing, China
9/2000 - 6/2004	**BS in Meteorology**, Nanjing University, Nanjing, China

Experience:

1/2010 - 6/2011 Research Internee, NASA Ames Ecological Forecasting Lab
Currently working on measurement of forest aboveground biomass based on data from both active and passtive remote sensors, such as Geoscience Laser Altimeter System (GLAS) LiDAR, Land, Vegetation, and Ice Sensor (LVIS), and Landsat.

6/2008 - 8/2008 Research Internee , NASA Science Mission Directorate Applied Sciences training and development program
The internship focused on regional estimation of evapotranspiration from remote sensing derived vegetation index and meteorological records. This work also revealed the regional water cycling for water management, and predicted the future water demand under given climate change scenarios.

8/2007 - 6/2012 Research Assistant, Utah State University
Building the theoretical model to link the remotely sensed NPP, biomass, and disturbance regime together, and solving the stand age for global primary forests.
Assembling ground phenology records from over 20 networks.

9/2004 - 6/2007 Research Assistant, Utah State University
Modeling the phenology of grassland in Inner Mongolia coupled water balance, and simulating phenology changes under regional climate scenario generated by RegCM3 together.

12/2003 - 6/2004 Research Assistant, Nanjing University
Retrieval past 2000 years land cover change and vegetation indices in China.

Awards:

8/2007 - 8/2012	Jesse E. Quinney Fellowship
6/2008 - 8/2008	NASA DEVELOP fellowship
4/2003	Excellent Research Paper Award of Atmospheric Science, Nanjing University

Publication:

G. Zhang et al. A Simple Parametric Estimation of Live Forest Aboveground Biomass in California Using Satellite Derived Metrics of Canopy Height and Leaf Area Index (Submitted to GRL)

S. Ganguly, R.R. Nemani, **G. Zhang**, Y. Knyazikhin, C Milesi, A Michaelis, H. Hashimoto, W. Wang, P. Votava, A. Samanta, F. Melton, J. Dungan, E. Vermote, F. Gao, R.B. Myneni, Generating Global Leaf Area Index from Landsat: Algorithm Formulation and Demonstration, 2012, Remote Sensing of Environment, Landsat Legacy Special Issue

G. Zhang, K. Lowry, R. Nemani, J. W. Skiles, and C. Schmidt.□Modeling Current and Future Water Use in Utah with Nasa's Terrestrial Observation and Prediction System, 2010, ASPRS Proceedings Vol. 30.

Gentile, A., **G. Zhang**, L. Pierce, G. Ciraolo, and G. La Loggia. Analysis of the energetic flows through the SEBAL application to the assemsment of the Actual Evapotraspiration in a Napa Valley Vineyard California , 2009, Proceeding of 12° AIAM Conference, , 12:15–17.

White, M. A., K. M. de Beurs, K. Didan, D. W. Inouye, A. D. Richardson, O. P. Jensen, J. O'KEEFE, **G. Zhang**, R. R. Nemani, and W. J. D. van Leeuwen. "Intercomparison, Interpretation, and Assessment of Spring Phenology in North America Estimated from Remote Sensing for 1982–2006." Global Change Biology 15, no. 10 (2009): 2335–2359.

A. Gentile, L. Pierce, G. Ciraolo, **G. Zhang**, G. La Loggia, R. Nemani, Comparison between energy balance and mass balance models for actual evapotranspiration assessment, 2009, Proc. SPIE, Vol. 7472, 747212

Presentation:

G. Zhang, S. Ganguly, R. R. Nemani, S. Saatchi, Y. Yu, A. Fore, W. Wang, P. Votava, A. Michaelis, R. B. Myneni, 2011 Estimation of Aboveground Biomass at a High Spatial Resolution Using an Extensive Data Record of Satellite Derived Metrics: A Case Study with California, NASA Carbon Cycle and Ecosystem Joint Science Workshp, 2011

G. Zhang, S. Ganguly, R. R. Nemani, P. Atkison, Jadu Dash, J. Chockalingam, J. L. Dungan, 2011 Validation of the Land Cover Dynamics Product from MODIS with Field Data and MERIS Derived Phenology Estimates over North America, AGU Fall Meeting, 2011

G. Zhang, and M. A. White, The Relationship Between Seasonal Pattern of Photosynthetic Parameters and Plant Phenology, AGU Fall Meeting, 2009.